IDVQ : l'indice d'impact de la douleur au cou et aux membres supérieurs sur la vie quotidienne

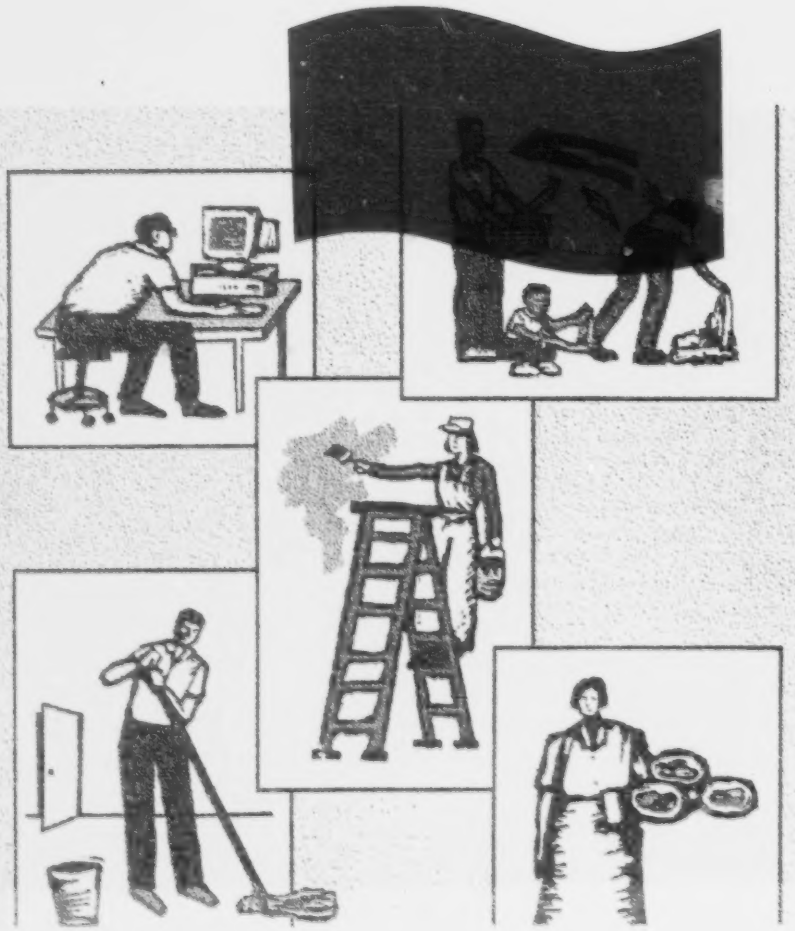
Développement et validation d'une nouvelle mesure de l'état de santé pour les travailleurs atteints de lésions musculo-squelettiques du cou et des membres supérieurs

Susan Stock

ÉTUDES ET RECHERCHES

R-355

RAPPORT



irsst

Solidement implanté au Québec depuis 1980, l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST) est un organisme de recherche scientifique reconnu internationalement pour la qualité de ses travaux.

NOS RECHERCHES
travaillent pour vous !

MISSION

- Contribuer, par la recherche, à la prévention des accidents du travail et des maladies professionnelles ainsi qu'à la réadaptation des travailleurs qui en sont victimes.
- Offrir les services de laboratoires et l'expertise nécessaires à l'action du réseau public de prévention en santé et en sécurité du travail.
- Assurer la diffusion des connaissances, jouer un rôle de référence scientifique et d'expert.

Doté d'un conseil d'administration paritaire où siègent en nombre égal des représentants des employeurs et des travailleurs, l'IRSST est financé par la Commission de la santé et de la sécurité du travail.

POUR EN SAVOIR PLUS...

Visitez notre site Web ! Vous y trouverez une information complète et à jour. De plus, toutes les publications éditées par l'IRSST peuvent être téléchargées gratuitement.
www.irsst.qc.ca

Pour connaître l'actualité de la recherche menée ou financée par l'IRSST, abonnez-vous gratuitement au magazine *Prévention au travail*, publié conjointement par l'Institut et la CSST.
Abonnement : 1-877-221-7046

Dépôt légal
Bibliothèque nationale du Québec
2003
ISBN : 2-551-22428-4
ISSN : 0820-8395

IRSST - Direction des communications
505, boul. De Maisonneuve Ouest
Montréal (Québec)
H3A 3C2
Téléphone : (514) 288-1551
Télécopieur : (514) 288-7636
publications@irsst.qc.ca
www.irsst.qc.ca
© Institut de recherche Robert-Sauvé en santé et en sécurité du travail, décembre 2003.

105-00110/

ÉTUDES ET RECHERCHES

IDVQ : l'indice d'impact de la douleur au cou et aux membres supérieurs sur la vie quotidienne

Développement et validation d'une nouvelle mesure de l'état de santé pour les travailleurs atteints de lésions musculo-squelettiques du cou et des membres supérieurs

RAPPORT

Susan Stock[1,2,3], Patrick Loisel[4], Marie-José Durand[4], David Streiner[5], Peter Tugwell[6], Rhoda Reardon[7], Jacques Lemaire[8], Micheline Boucher[8], Susan Darzins[9], Peter Dilworth[10] et Nathaly Gaudreault[11]

[1]Direction de la santé publique de Montréal-Centre
[2]Institut national de santé publique du Québec
[3]Université McGill
[4]PRÉVICAP, Université de Sherbrooke
[5]Université de Toronto
[6]Université d'Ottawa
[7]Institut de recherche sur le travail et la santé (Ontario)
[8]Centre des recherches cliniques de l'Hôpital Charles LeMoyne
[9]Université de Melbourne, Australie
[10]Université McMaster
[11]Université de Montréal

Cliquez recherche
www.irsst.qc.ca

Cette publication est disponible en version PDF sur le site Web de l'IRSST.

Cette étude a été financée par l'IRSST. Les conclusions et recommandations sont celles des auteurs.

CONFORMÉMENT AUX POLITIQUES DE L'IRSST

Les résultats des travaux de recherche publiés dans ce document ont fait l'objet d'une évaluation par des pairs.

REMERCIEMENTS

Nous aimerions d'abord remercier de leur collaboration inestimable et de leur disponibilité tous les physiothérapeutes et ergothérapeutes qui nous ont accueilli dans leurs cliniques ainsi que tous les travailleurs et travailleuses atteints de lésions musculo-squelettiques de ces cliniques qui ont participé aux divers volets de ce projet. Ces travailleurs ont contribué énormément à notre compréhension de leur vécu lorsqu'ils sont atteints de lésions musculo-squelettiques. Les cliniques suivantes ont participé à l'étude québécoise : quatre cliniques de physiothérapie de monsieur Pierre Aswad (sises à Longueuil, Brossard, Greenfield Park et St-Hubert), la clinique de Longueuil de monsieur Gilles Messier, les cliniques de Kinatex, de l'Institut du dos de Montréal, Esprit-Sport du secteur l'Acadie, celles de l'Institut de physiatrie du Québec et la clinique Centre-Ville de Montréal, ainsi que la Clinique Physiotech de Jean Luc Simard située à Pointe-aux-Trembles et les deux cliniques de physiothérapie de monsieur Richard Messier à Granby et à Cowansville en Estrie.

Nos remerciements s'adressent également aux deux équipes d'agents de recherche et d'intervieweurs qui ont participé au projet. Leur contribution a été essentielle à la réalisation de ce projet. Susan Darzins a assuré la coordination en Ontario et Micheline Boucher au Québec. Nous remercions Suzanne Deguire pour la coordination qu'elle a faite pendant une période de congé de Micheline Boucher. Nos intervieweurs au Québec incluaient Josette Dupuis, Céline Plessis-Bélair, Lucie Blain et Charlotte Fournier.

Nous tenons à remercier Dr Michael Abrahamowicz du Département d'épidémiologie, de biostatistique et de santé au travail de l'Université McGill pour ses conseils méthodologiques qui ont influencé les analyses réalisées.

Nos remerciements vont également à Carole Bellazzi de l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail pour son soutien général au projet et son assistance précieuse pour la révision du rapport ainsi qu'à Maryse Arpin pour son support de secrétariat y compris le travail de mise en page du rapport. Nous mentionnons également la contribution de Natalie Kishchuk à la rédaction d'une version préliminaire du rapport.

Finalement nous aimerions remercier tous les organismes subventionnaires qui, par leur soutien financier, ont permis la réalisation de ce projet. Le soutien financier de la Fondation Hamilton, du Ministère de la santé de l'Ontario et de l'Institut de recherche sur le travail et la santé d'Ontario (à l'époque appelé «Ontario Workers' Compensation Institute») a permis la réalisation des trois premières phases du projet soit l'identification des items potentiels, le raffinement des questions et l'étude de sélection des items en anglais dans une population ontarienne. La subvention de l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail a permis la réalisation des autres phases de ce projet incluant la traduction en français des items potentiels, la sélection des items en français dans une population québécoise et les analyses statistiques.

RÉSUMÉ

Ce document décrit le développement d'un nouvel outil mesurant l'état de santé spécifique aux lésions musculo-squelettiques du cou et des membres supérieurs. Il a été conçu de façon à pouvoir mesurer l'évolution des changements chez chaque sujet au cours du temps dans des études d'intervention et de discerner la sévérité des lésions des sujets dans les études étiologiques, pronostiques ou de prévalence. L'outil a été conçu pour s'appliquer aussi bien aux travailleuses et travailleurs francophones qu'anglophones du Québec et du Canada, et pour s'utiliser de façon pratique et facile tant en milieux cliniques tant qu'en recherche.

Le développement de l'outil a inclut : la génération et le prétest des questions potentielles; la réalisation d'études pilotes de fiabilité inter-juges des questions potentielles et des examens physiques; une étude pour la sélection préliminaire des items en anglais; la traduction en français des questions potentielles; une étude de sélection des items en français ; un comparaison des résultats des deux études, la sélection de 20 questions définitives et l'évaluation de la fiabilité, la validité et la sensibilité au changement de ces 20 questions dans les deux populations. Les questions potentielles et plusieurs mesures de validation ont été administrées, avant et après une série de traitements, à des travailleurs souffrant de lésions musculo-squelettiques des tissus mous des membres supérieurs ou du cou de deux populations, une anglophone et ontarienne et l'autre francophone et québécoise, recrutés à travers des centres de physiothérapie et d'ergothérapie en Ontario et au Québec. La sélection des items a été réalisée par l'application d'une combinaison de critères psychométriques et clinimétriques.

La version définitive de l'outil, intitulé « Indice d'impact de la douleur au cou et aux membres supérieurs sur la vie quotidienne » (IDVQ-20), comprend 20 questions évaluant les impacts des lésions musculo-squelettiques du cou et des membres supérieurs sur le travail, les activités physiques, le sommeil, les dimensions psychosociales ainsi que les effets iatrogènes chez les personnes atteintes. L'appellation en anglais de cette mesure est « Neck and Upper Limb Index » (NULI-20).

Les résultats ont démontré que l'IDVQ-20 possède une forte fiabilité test-retest ($CIC_{(1,1)}$ de 0,88 à l'étude anglophone et 0,83-0,85 à l'étude francophone) et cohérence interne (Alpha de Cronbach de 0,90 pour les sujets de l'étude anglophone et de 0,92 à l'étude francophone) aussi bien qu'une sensibilité au changement très robuste (moyenne des réponses normalisées de 1,48 pour la version anglaise et de 1,74 pour la version française). Cet outil présente également une bonne validité convergente selon les corrélations à plusieurs mesures objectives et subjectives d'examens physiques, de douleur et de mesure générique de l'état de santé. Dans des études de validation de construit, il a aussi été capable de discriminer de façon significative entre les personnes absentes du travail et celles qui travaillaient.

L'IDVQ-20 fournit donc une évaluation fiable et valide des impacts spécifiquement liés aux lésions musculo-squelettiques du cou et des membres supérieurs, dans les principales dimensions de la vie touchées par ces lésions. Il se présente sous un format court et facile à compléter.

TABLE DES MATIÈRES

REMERCIEMENTS I

RÉSUMÉ III

TABLE DES MATIÈRES V

LISTE DES TABLEAUX VII

LISTE DES FIGURES VIII

LISTE DES ANNEXES IX

1. PROBLÉMATIQUE 1

2. MÉTHODES 5

2.1 PHASE 1 : IDENTIFICATION DES ITEMS POTENTIELS 6

2.2 PHASE 2 : RAFFINEMENT DES QUESTIONS 7

2.2.1 Formulation des questions 7

2.2.2 Échelles de réponse et calcul des scores 7

2.2.3 Révision de langage pour assurer la compréhension par des personnes de faible scolarité 8

2.2.4 Prétest 8

2.2.5 Formation des intervieweurs et test de fiabilité inter-juges 8

2.3 PHASE 3 : ÉTUDE DE SÉLECTION DES ITEMS EN ANGLAIS 9

2.3.1 Population à l'étude 9

2.3.2 Procédure de l'étude 10

2.3.3 Évaluation et sélection des items de l'outil en anglais 12

2.3.3.1 Applicabilité des questions 12

2.3.3.2 Reproductibilité (fiabilité test-retest) 12

2.3.3.3 Cohérence interne 12

2.3.3.4 Sensibilité au changement 13

2.3.3.5 Capacité discriminatoire 13

2.3.3.6 Sélection parmi les questions restantes 14

2.4 PHASE 4 : TRADUCTION EN FRANÇAIS DES ITEMS POTENTIELS ET DES QUESTIONNAIRES DE VALIDATION. 15

2.5 PHASE 5 : ÉTUDE DE SÉLECTION DES ITEMS EN FRANÇAIS 16

2.5.1 Formation des intervieweurs et tests de fiabilité inter-juges 16

2.5.2 Population à l'étude 16

2.5.3 Procédures de l'étude 17

2.5.4 Évaluation et sélection des items de l'outil en français 17

2.6 PHASE 6 : COMPARAISON DES RÉSULTATS DES ÉTUDES ANGLOPHONES ET FRANCOPHONES 17

2.7 PHASE 7 : SÉLECTION DES 20 ITEMS APPROPRIÉS AUX DEUX POPULATIONS 19

2.8 PHASE 8 : TESTS DE FIABILITÉ, DE VALIDITÉ ET DE SENSIBILITÉ AU CHANGEMENT DE LA VERSION FINALE 20

2.8.1 Fiabilité 20

2.8.1.1 Cohérence interne 20

2.8.1.2 Fiabilité test-retest 20

2.8.2 Validité 21

2.8.2.1 Validité apparente 21

2.8.2.2 Validité convergente 21

2.8.2.3 Validité de construit 22

2.8.3 Sensibilité au changement 22

3. RÉSULTATS 25
3.1 Identification des items potentiels 25
3.2 Fiabilité inter-juges des items potentiels 25
3.2.1 Étude anglophone de fiabilité inter-juges *25*
3.2.2 Étude francophone de fiabilité inter-juges *26*
3.3 Sélection des items 27
3.3.1 Description des populations à l'étude *27*
3.3.1.1 Population de l'étude anglophone 27
3.3.1.2 Population de l'étude francophone 27
3.3.2 Élimination et sélection des items de l'étude anglophone *29*
3.3.3 Élimination et sélection des items de l'étude francophone *31*
3.4 Importance relative des domaines 35
3.5 Comparaison des résultats des études anglophones et francophones 37
3.5.1 Comparaison des scores moyens de la question globale de chaque domaine *37*
3.5.2 Comparaison du classement des domaines *39*
3.5.3 Interprétation des comparaisons *40*
3.6 Sélection définitive des 20 items appropriés aux deux populations 41
3.6.1 Analyses factorielles *42*
3.7 Évaluation de la fiabilité de l'IDVQ-20 43
3.7.1 Cohérence interne *43*
3.7.2 Fiabilité test-retest *43*
3.8 Validité de l'IDVQ-20 44
3.8.1 Validité convergente *44*
3.8.2 Validité de construit *45*
3.9 Sensibilité au changement 47
4. DISCUSSION 49
5. CONCLUSION 55
6. APPLICABILITÉ DES RÉSULTATS ET RETOMBÉES ÉVENTUELLES 57
7. LISTE DES ARTICLES SCIENTIFIQUES PRODUITS DANS LE CADRE DE CETTE SUBVENTION 59
RÉFÉRENCES 61
ANNEXES

LISTE DES TABLEAUX

Tableau 1. : Comparaison des populations à l'étude.

Tableau 2. : Résultats de la démarche d'élimination et de sélection des 35 items reformulés de l'IDVQ-35.

Tableau 3. : Relation entre les scores moyens de la question globale de chaque domaine et le statut du groupe d'étude (anglophone vs francophone) en contrôlant pour l'âge, le sexe, le revenu et la durée des symptômes.

Tableau 4. : Comparaison du rang moyen de chaque domaine entre l'étude francophone et l'étude anglophone à l'aide des tests de « Wilcoxon Rank Sum ».

Tableau 5. : Relation entre les scores moyens de classement de chaque domaine et le statut du groupe d'étude (anglophone vs francophone) en contrôlant pour l'âge, le sexe, le revenu et la durée des symptômes.

Tableau 6. : Fiabilité et sensibilité au changement de l'IDVQ-20.

Tableau 7. : Comparaisons des coefficients de corrélation de Pearson entre l'IDVQ-20 et les autres mesures.

Tableau 8. : Comparaison du score moyen à la fin des traitements, des 20 questions de l'IDVQ-20 et de la sous échelle sur le travail, entre les répondants en arrêt de travail à la fin des traitements et ceux qui étaient au travail.

Tableau 9. : Comparaison du score moyen à la fin des traitements, des 20 questions de l'IDVQ-20 et de la sous échelle sur le travail, chez les répondants en arrêt de travail au début des traitements, entre ceux qui sont retournés au travail avant la fin des traitements et ceux qui sont restés en arrêt de travail.

Tableau 10. : Comparaison des MRN de l'IDVQ et les autres mesures.

LISTE DES FIGURES

Figure 1. : Comparaison du score moyen de la question globale de chaque domaine en fonction du groupe d'étude (anglophone vs francophone).

Figure 2. : Comparaison du rang moyen de chaque domaine en fonction du groupe d'étude (anglophone vs francophone).

LISTE DES ANNEXES

Annexe A : Résultats de fiabilité inter-juges (ICC) de chaque item selon le domaine

Annexe B : L'IDVQ-35

Annexe C : Résultats des critères d'élimination et de sélection pour les 20 items finals sélectionnés à l'étude anglophone

Annexe D : Résultats des critères d'élimination et de sélection pour les 20 items finals sélectionnés à l'étude francophone (formulation originale)

Annexe E : IDVQ-20 Version finale en français

Annexe F : NULI-20 Version finale en anglais

1. PROBLÉMATIQUE

Les lésions musculo-squelettiques (LMS) du cou et des membres supérieurs liées au travail sont une cause importante d'incapacité dans la population et sont très coûteuses pour la société (National Research Council 2001, Arcand et al 2000, Buckle et Devereux 1999). Malgré la grande prévalence de ces lésions, il n'y a pas de consensus sur les interventions les plus appropriées pour leur traitement ou leur prévention. Des études bien conçues et rigoureuses sont requises pour évaluer l'efficacité des interventions proposées. Une des conditions préalables pour réaliser de telles études est la disponibilité d'une mesure de résultats standardisée, fiable et valide appropriée à ces lésions. Les mesures de statut fonctionnel ou de l'état de santé peuvent servir à cette fonction dans une grande variété de recherches et d'applications cliniques.

Les mesures de l'état de santé («*health status measures*») prennent la forme de questionnaires adressés directement aux répondants qui quantifient les symptômes, le fonctionnement physique, les sentiments et/ou le comportement. Leur portée peut varier considérablement (Bell et al. 1990, Bergner 1989, Cohen et Marino 2000, Feinstein et al. 1986, Gill et Feinstein 1994, Greenfield et Nelson 1992, Jette 1984, Naughton et Shumaker 1997, Testa et Simonson 1996, Ware 1984, 1995). Ces mesures permettent de décrire l'impact des maladies ou des problèmes de santé sur plusieurs dimensions de la vie en se basant sur l'expérience du patient plutôt que sur l'interprétation d'un professionnel de la santé. Les mesures de l'état de santé peuvent être des outils simples axés sur le fonctionnement physique de l'individu qui sont généralement définis en fonction des activités de soins personnels et de mobilité. Ils peuvent être des outils plus larges qui mesurent plusieurs dimensions du fonctionnement physique de la personne (ex. : tâches domestiques et familiales, travail, activités de loisir, soins personnels, transport, etc.). Finalement, les mesures de l'état de santé peuvent aussi inclure des outils de « qualité de vie liée à la santé » qui mesurent non seulement le fonctionnement physique de la personne mais aussi son fonctionnement psychologique, social et de rôle.

Les mesures de l'état de santé peuvent être utilisées dans des définitions de cas distinguant le niveau de sévérité du problème de santé étudié dans des études étiologiques, de pronostic ou de prévalence. Elles peuvent aussi servir à mesurer le changement de l'état de santé chez un

individu au cours du temps dans des études d'évaluation de l'efficacité d'interventions de prévention primaire, de traitement ou de réadaptation et de prévention d'incapacité à long terme aussi bien que dans des études d'évaluation de l'impact des services de santé (Deyo et Carter 1992, Greenfield et Nelson 1992, Fries 1983, Patrick 1997). Les mesures de l'état de santé sont de plus en plus reconnues comme mesures d'impact extrêmement pratiques, susceptibles de capter le fardeau des maladies de façon plus exacte que les mesures de déficiences physiologiques (Deyo 1988, Feinstein et al. 1986). Elles sont aussi précieuses dans le contexte clinique, puisqu'elles fournissent un outil d'évaluation rapide et fiable pour suivre le progrès clinique des patients ou pour dépister les problèmes de fonctionnement physiques et psychosociaux (Deyo et Carter 1992, Greenfield et Nelson 1992, Guillemin 2000). Elles pourraient également jouer un double rôle dans les milieux du travail d'abord, en surveillance des travailleurs pour déceler la gravité des lésions musculo-squelettiques et ensuite, en évaluant les programmes d'intervention mis en place pour réduire l'incidence ou la prévalence des LMS.

Les mesures d'état de santé sont particulièrement pertinentes dans le contexte d'étude des lésions musculo-squelettiques où des outils valides et fiables spécifiques aux maux de dos et à certaines maladies rhumatologiques ont déjà été développés avec succès. De telles mesures ont démontré qu'elles sont égales ou meilleures – au point de vue reproductibilité et sensibilité au changement – que les tests diagnostics cliniques soi-disant « objectifs » (ex. : test d'examen physique, radiographies, tests de conduction nerveuse) (Ahlmen et al 1988, Deyo 1988, Fries 1983, Katz et al. 1994, Levine et al 1993, Wilson et Cleary 1995). Des mesures de l'état de santé spécifiques aux lésions musculo-squelettiques se sont également montrées plus sensibles au changement que les mesures de l'état de santé génériques (Amadeo et al 1996, Bessette et al 1998, Bombardier et al 1995, Patrick et Deyo 1989, Stucki et al 1995).

Ce document décrit le développement d'une nouvelle mesure de l'état de santé spécifique aux lésions musculo-squelettiques du cou et des membres supérieurs qui mesure l'impact de ces lésions sur les activités physiques, le travail, le sommeil ainsi que sur les dimensions de l'expérience psychologique et sociale.

Ce projet avait pour objectif de développer une mesure fiable et valide de qualité de vie liée à la santé :

- spécifique aux lésions musculo-squelettiques du cou et des membres supérieurs,
- capable de mesurer l'évolution des changements de l'état de santé musculo-squelettique chez les mêmes sujets au cours du temps dans des études d'intervention,
- capable de distinguer entre les sujets (c'est-à-dire évaluer la gravité des lésions) en études d'étiologie, de pronostic ou de prévalence,
- applicable à la population francophone aussi bien qu'anglophone du Canada,
- pratique et facile à utiliser dans les milieux cliniques.

Le but était de répondre à tous les critères méthodologiques identifiés par Tugwell et Bombardier (1982), par Kirshner et Guyatt (1985) et par Guyatt et al (1996) pour juger des mesures de l'état de santé.

Pour nous assurer que le ou les mesures développées sont appropriées pour les deux principaux groupes culturels et linguistiques du Canada (anglais et français), nous avons mené des études indépendantes et parallèles pour la sélection des questions et les validations subséquentes dans des populations anglophones et francophones distinctes. Une étude a été menée dans une population anglophone de l'Ontario et l'autre dans une population francophone du Québec. Les protocoles des deux études étaient semblables. Cela contraste avec la procédure habituelle qui consiste à développer une mesure dans un seul groupe culturel linguistique, de traduire le produit définitif pour l'utiliser dans d'autres groupes culturels et linguistiques puis de vérifier sa fiabilité ou sa validité.

2. MÉTHODES

Cet indice a été élaboré selon les méthodes recommandées par Streiner et Norman (1995), par Juniper et al (1996) et par Guyatt et al (1986, 1996). Il comprend les étapes suivantes :

Phase 1 Identification des items potentiels

Phase 2 Raffinement des questions : pré-test et étude de fiabilité inter-juges

Phase 3 Étude de sélection des items en anglais

Phase 4 Traduction en français des items potentiels et des questionnaires de validation

Phase 5 Étude de sélection des items en français

Phase 6 Comparaison des résultats anglais et français

Phase 7 Sélection des 20 items appropriés aux deux populations

Phase 8 Tests de fiabilité, de validité et de sensibilité au changement de la version finale.

2.1 Phase 1 : Identification des items potentiels

La stratégie utilisée pour identifier les domaines ou les dimensions de la vie touchés par les lésions musculo-squelettiques du cou et des membres supérieurs et un ensemble d'items potentiels reflétant ces domaines a été décrite antérieurement (Stock 1991, Stock et al 1996). Le terme «items» réfère aux questions individuelles incluses dans les questionnaires sur l'état de santé existants ou les questions potentielles identifiées dans cette étude et par la suite, testées et sélectionnées ou éliminées lors de la mise au point de l'indice final. Le processus d'identification des items potentiels a inclus les étapes suivantes effectuées en Ontario :

(1) une revue de la littérature scientifique pertinente en rhumatologie, en réadaptation, en orthopédie et concernant la problématique des lombalgies;

(2) un recensement des items inclus dans des mesures de l'état de santé génériques ou des mesures de statut fonctionnel spécifiques aux LMS (créés par des chercheurs ou des cliniciens en rhumatologie, en réadaptation, en orthopédie ou en santé communautaire). Ces mesures ont été identifiées par la revue de la littérature et par la consultation d'experts dans le domaine;

(3) l'expérience clinique des chercheurs.

(4) des entrevues avec 33 patients-travailleurs souffrant de lésions musculo-squelettiques du cou ou des membres supérieurs réalisées dans des cliniques de santé au travail situées dans cinq milieux différents (une grande usine sidérurgique, une manufacture de taille moyenne, un hôpital pour soins chroniques, une clinique communautaire de santé au travail et une clinique universitaire de santé au travail) pour identifier l'impact de leurs lésions musculo-squelettiques sur les diverses dimensions de leur vie;

(5) un sondage par questionnaire réalisé auprès de 30 cliniciens (y compris physiothérapeutes, ergothérapeutes, rhumatologues, physiatres, chirurgiens orthopédistes et médecins du travail) qui traitent des travailleurs souffrant de lésions musculo-squelettiques touchant le cou et les membres supérieurs;

2.2 Phase 2 : Raffinement des questions

2.2.1 Formulation des questions

Les questions potentielles générées par la stratégie mentionnée précédemment ont été catégorisées en douze domaines (ou dimensions) de vie susceptibles d'êtres touchés par ces lésions. Les questions redondantes ont été éliminées et un questionnaire comportant un groupe de questions potentielles pour chaque domaine a été élaboré. Chaque domaine comptait entre sept et trente questions spécifiques et une question générale concernant l'impact global de la lésion dans ce domaine de la vie. En ce qui a trait aux domaines concernant des activités (ex. : travail, responsabilités domestiques et familiales, le transport et l'usage de la voiture, les loisirs, les activités sociales, etc.), le degré d'applicabilité de chaque item a été évalué par le biais d'une question de dépistage demandant au sujet si l'activité mentionnée s'appliquait à sa situation. Le répondant devait ensuite évaluer le degré de difficulté d'exécution des activités identifiées comme applicables ou pertinentes. L'applicabilité des items relatifs à la vie sexuelle était évaluée, quant à elle, par une seule question de dépistage visant à identifier les sujets sexuellement inactifs pour des raisons autres que leurs problèmes de cou ou de membres supérieurs.

2.2.2 Échelles de réponse et calcul des scores

Toutes les questions ont été présentées sur une échelle numérique de sept points, comportant également des descriptions verbales, pour en maximiser la fiabilité. Le choix du type d'échelle et du nombre de points de l'échelle a été fondé sur l'évidence fournie par la littérature psychométrique suggérant que la fiabilité d'une échelle s'élève rapidement jusqu'à un seuil de sept points, puis plus lentement jusqu'à un seuil de onze points (Nunnally 1978, Nunnally et Wilson 1975, Nishisato et Torii 1970). Bien qu'aucun avis unanime n'appuie ce point, Streiner et Norman (1995) ont recommandé le recours à des échelles d'au moins cinq à sept points. Des descriptions verbales ont été incluses sous chaque chiffre puisque l'ajout des descriptions verbales augmente la fiabilité de l'échelle (Streiner 1985).

Pour les questions concernant les activités, l'échelle de sept points était exprimée en degré de difficulté. Pour les questions concernant les autres domaines (ex : humeur; estime de soi, sommeil, aspects iatrogéniques, etc.) les échelles de sept points retenues exprimaient soit la fréquence de l'énoncé, soit le degré d'accord ou désaccord avec celui-ci.

Le score moyen de chaque domaine a été calculé par l'addition des scores de chaque question applicable divisée par le nombre des questions applicables. Plus le score est élevé, plus l'incapacité associée à la lésion est importante. Un score plus bas indique une incapacité moins importante.

2.2.3 Révision de langage pour assurer la compréhension par des personnes de faible scolarité

Le questionnaire qui a découlé de l'étape précédente a ensuite été soumis à un rédacteur d'alphabétisation pour assurer que des personnes possédant une 6^{e} année de scolarité puissent bien comprendre le niveau de langage utilisé.

2.2.4 Prétest

Pour identifier les questions obscures, ambiguës, difficiles à comprendre ou inappropriées, le questionnaire a été prétesté auprès de dix clients anglophones souffrant de lésions au cou ou aux membres supérieurs. Ces clients se sont présentés à la clinique de réadaptation professionnelle d'un hôpital universitaire exécutant des évaluations de capacité fonctionnelle chez les travailleurs blessés. Des protocoles détaillés pour les examens physiques ont aussi été élaborés conjointement aux questionnaires.

2.2.5 Formation des intervieweurs et test de fiabilité inter-juges

Des personnes ayant une formation de professionnel de la santé ont été recrutées comme intervieweurs et formées pour administrer le questionnaire et suivre les protocoles standardisés des examens physiques utilisés dans l'étude de validation. Des séances de formation de trois à cinq jours ont été tenues afin que les intervieweurs se familiarisent avec le contenu des questions,

l'utilisation des échelles et qu'on leur enseigne une technique standardisée appropriée d'entrevues et de procédures d'examens physiques. Les intervieweurs ont été formés pour poser des questions d'une manière non dirigée et impartiale et adopter une attitude interpersonnelle neutre. Les séances de formation comprenaient une rétroaction sur les entrevues enregistrées et des jeux de rôle d'entrevues avec des cas susceptibles d'être difficiles.

Pour assurer la cohérence entre les intervieweurs, une étude de fiabilité inter-juges a été menée en utilisant le questionnaire révisé des questions potentielles. En Ontario, trente-huit patients-travailleurs souffrant de LMS du cou ou des membres supérieurs de quatre milieux cliniques ont participé à cette étude précédant l'étude de sélection des questions. Deux intervieweurs ont interrogé et examiné chaque patient le même jour, à intervalles de deux à trois heures. Après la seconde entrevue, la rétroaction des répondants a été demandée afin d'identifier plus précisément toute question ambiguë ou difficile à comprendre.

La fiabilité inter-juges a été mesurée avec des coefficients de corrélation intraclasse ($CIC_{(1,1)}$) (Shrout et Fleiss 1979) calculés pour la moyenne des scores des questions dans chaque domaine et de chaque question individuelle. Les questions démontrant une fiabilité inter-juges de moins de 0,7 ont été identifiées et la source de la difficulté a été déterminée à l'aide de discussions avec les intervieweurs. Suite à ces réflexions, certaines questions ont été reformulées.

2.3 Phase 3 : Étude de sélection des items en anglais

L'étude ontarienne a reçu l'approbation du comité d'éthique du Research Ethics Committee de l'hôpital universitaire Chedoke-McMaster.

2.3.1 Population à l'étude

La population ontarienne à l'étude comprenait 119 travailleurs anglophones qui ont subi des lésions musculo-squelettiques des tissus mous du membre supérieur ou du cou. Ils ont été recrutés dans huit cliniques communautaires de physiothérapie à Hamilton en Ontario de février 1993 à juin 1994. Au moment de l'étude, les sujets devaient occuper un emploi rémunéré et travailler ou être absent du travail à cause de leur problème musculo-squelettique. Ils devaient

aussi posséder une bonne compréhension de l'anglais et être capables de communiquer dans cette langue. Les cliniques de physiothérapie participant à l'étude ont fourni des services de réadaptation subventionnés par la Commission des accidents du travail de l'Ontario (Ontario Workers' Compensation Board (WCB)) aux patients souffrant de lésions musculo-squelettiques professionnelles ainsi qu'aux patients ayant subi des lésions non liées au travail. Le programme de réadaptation exigé par le WCB et offert aux travailleurs blessés comprenait l'application de diverses modalités physiques pour le soulagement de la douleur, un programme d'éducation et un programme d'exercices progressifs d'étirement et de renforcement musculaire et d'aérobie. (Kaegi 1990, Mitchell et Carmen 1990, VanShoor et Wright 1989). Ce programme durait jusqu'à six semaines à raison de cinq jours par semaine, passant graduellement d'une durée de une heure à trois ou quatre heures par jour. Le programme des patients qui se rétablissaient plus rapidement a été plus court. Les patients non indemnisés par le WCB ont reçu des services similaires mais moins intensifs (ex : séances de 20 à 45 minutes, 2 à 3 fois par semaine). Des informations démographiques de base ont été obtenues de toutes les personnes qui n'ont pas accepté de participer à l'étude.

2.3.2 Procédure de l'étude

Les renseignements sociodémographiques recueillis incluaient l'âge, le sexe, l'éducation, le revenu familial, le niveau de compréhension de l'anglais, le statut d'immigration, la langue parlée à la maison, la situation relative à l'emploi, la profession, l'industrie et la date du début des symptômes. Le questionnaire, comprenant environ 175 questions potentielles, a été administré à tous les travailleurs consentants inscrits au programme de traitement aux cliniques de physiothérapie et souffrants de LMS au cou et aux membres supérieurs. Il a été administré au début du traitement, deux à sept jours plus tard (pour vérifier sa fiabilité «test-retest»), et le dernier jour du programme de traitement. Lors du premier jour de l'administration du questionnaire et au moment des entrevues post-traitement, les questionnaires et les examens physiques additionnels suivants ont été administrés :

(i) une question unique mesurant l'évaluation globale du sujet concernant l'impact de sa lésion musculo-squelettique sur sa qualité de vie ;

(ii) une question analogue mesurant l'évaluation globale du physiothérapeute traitant concernant l'impact de la lésion musculo-squelettique sur la qualité de vie de son client ;

(iii) le questionnaire *Sickness Impact Profile* (S.I.P.) (Bergner et al 1981) ;

(iv) une mesure de l'intensité de la douleur sur une échelle chiffrée de 0 à 10, sous forme de thermomètre permettant de mesurer l'intensité de la douleur à chaque partie du corps atteinte identifiée par le répondant sur un schéma corporel des douleurs (ex. : cou, épaules, coudes, avant-bras, poignets, mains) ;

(v) une mesure de la force de préhension de la main atteinte à l'aide d'un dynamomètre Jamar, selon un protocole écrit ;

(vi) une mesure de l'abduction de l'épaule atteinte mesurée à l'aide d'un goniomètre, selon un protocole écrit ;

(vii) le test « main-nuque » de l'épaule atteinte (distance en centimètres entre la charnière cervico-dorsale, c'est-à-dire la protubérance à la nuque au niveau de la 7^{e} vertèbre cervicale et l'index lorsque l'épaule effectue une rotation interne et que la main tente de rejoindre le point le plus haut du dos, mesuré avec un ruban à mesurer) ;

(viii) des questions concernant la situation relative à l'emploi (ex. : travail régulier, travail léger ou modifié, arrêt de travail).

Au cours de l'entrevue initiale, les répondants ont aussi été priés de classer par ordre d'importance les douze domaines de vie selon l'importance relative de l'impact de leur lésion musculo-squelettique sur ces divers domaines.

La pertinence de recourir à une question unique pour mesurer l'évaluation globale du sujet ou du physiothérapeute concernant l'impact d'un problème de santé sur sa qualité de vie est décrite par Gill et Feinstein (1994).

En terme de déroulement, un intervieweur a administré le questionnaire comprenant les items potentiels et le questionnaire sociodémographique. Un autre intervieweur a administré les questionnaires associés aux mesures de validation soit le SIP, la question globale unique sur l'impact de la LMS sur la qualité de vie (posée au patient et au physiothérapeute traitant), et a effectué les examens physiques. Ce dernier né connaissait donc pas les réponses aux questions

posées par le premier intervieweur concernant les items potentiels. La deuxième administration du questionnaire pour vérifier la fiabilité test-retest des items potentiels a été limitée à 100 sujets.

2.3.3 Évaluation et sélection des items de l'outil en anglais

L'objectif de la réduction des items était d'identifier et de retirer les questions qui n'étaient pas applicables ou pertinentes à une proportion significative des répondants, dont la fiabilité test-retest était faible, dont la capacité de discriminer entre les sujets était limitée ou qui n'étaient pas sensibles au changement. En conséquence, les cinq critères retenus pour procéder à l'élimination et la sélection des questions étaient : l'applicabilité, la reproductibilité, la cohérence interne, la sensibilité au changement et la capacité discriminatoire des questions.

2.3.3.1 Applicabilité des questions

L'applicabilité des questions a été mesurée en évaluant le pourcentage des sujets qui ont déclaré la question pertinente pour eux lors de la question de dépistage mentionnée précédemment. Pour chaque item, ce pourcentage a été calculé pour l'ensemble de la population à l'étude et pour des sous-groupes selon le sexe et l'âge. Les questions qui n'ont pas été applicables à au moins 80 % de la population totale ont été exclues. Les questions qui n'ont pas été applicables à au moins 70 % des hommes et 70 % des femmes ont également été exclues.

2.3.3.2 Reproductibilité (fiabilité test-retest)

La reproductibilité de chaque question a été mesurée par le coefficient de corrélation de Pearson corrélant les points de la première administration de la question à la seconde, deux à sept jours plus tard. Les questions dont le coefficient de corrélation était de moins de 0,50 ont été exclues.

2.3.3.3 Cohérence interne

Pour chaque domaine, la cohérence interne de chaque question a été mesurée en calculant les coefficients de corrélation entre le score de la question et 1) le score moyen de l'ensemble des

questions dans le domaine sans cette question; et 2) le score de la question globale pour le domaine. Les questions dont l'une ou l'autre corrélation était de moins de 0,30 ont été exclues.

2.3.3.4 Sensibilité au changement

La sensibilité au changement de chaque question a été mesurée en calculant la corrélation entre le changement résiduel des scores (CRS) de la question avant et après les traitements et le changement résiduel des scores de la question globale du domaine avant et après les traitements. Les questions dont la corrélation entre ces changements résiduels des scores étaient de moins de 0,30 ont été exclues. Aussi, la corrélation entre le CRS de l'item et le CRS de la moyenne de la question unique globale sur l'impact de la lésion sur la qualité de vie posée au sujet et à son clinicien a été évaluée. Les questions dont la corrélation entre ces changements résiduels des scores étaient de moins de 0,30 ont été exclues. Il a été décidé d'évaluer le CRS comme mesure de changement au lieu du score absolu de changement afin d'éviter les problèmes statistiques de la régression à la moyenne associés au score absolu de changement (Nunnally 1975, Norman 1989, Francis et al 1991).

2.3.3.5 Capacité discriminatoire

Afin d'évaluer la capacité discriminatoire de chaque question, des histogrammes de la distribution des scores sur l'échelle de sept points ont été calculés pour chaque question et ces histogrammes ont été observés pour les signes d'asymétrie. Les questions dont la distribution était très asymétrique ont été exclues. En règle générale, nous avons exclu les questions dont le coefficient d'asymétrie («skewness») a été de plus de 2 écarts types de l'erreur type de cette statistique. Dans quelques cas, les observations des histogrammes nous ont convaincus que la distribution des réponses était acceptable malgré un coefficient d'asymétrie élevé. Cette situation est arrivée quand 25 à 30 % des répondants ont coté un score élevé sur une question et les histogrammes ont démontré une bonne distribution des 70 % autres répondants sur les autres points plus faibles de l'échelle de réponse. Nous avons jugé qu'il était pertinent d'inclure quelques questions dont un tiers des répondants indiquaient que cet item a eu un impact important sur eux.

De plus, le score de chaque question a été corrélé avec le score de la question mesurant l'évaluation globale du sujet concernant l'impact de sa lésion sur sa qualité de vie, la question analogue mesurant l'évaluation globale du physiothérapeute traitant concernant l'impact de la lésion sur la qualité de vie de son client et la moyenne de ces deux questions d'évaluation globale. Les questions qui ne faisaient pas l'objet d'une corrélation statistiquement significative ont été exclues.

2.3.3.6 Sélection parmi les questions restantes

Une fois les questions correspondant aux critères d'exclusion mentionnés ci-dessus éliminées, il s'est avéré nécessaire d'établir une priorité parmi les questions restantes. Cette étape a non seulement été effectuée au niveau des questions individuelles, mais aussi au niveau des domaines. Les résultats des scores moyens de la question globale pour chaque domaine et les résultats des classements des domaines ont été utilisés pour évaluer l'importance et l'impact relatifs attribués à chaque domaine par les sujets à l'étude. Ces données ont été calculées pour la population à l'étude et selon le sexe. Un comité de co-chercheurs (MJD, JL, PL, DS, SS, PT, MA) a étudié ces données et suite à une discussion consensuelle, le choix des domaines de vie prioritaires à retenir et le nombre de questions dans chaque domaine ont été déterminés.

La sélection parmi les questions restantes visait à choisir les éléments les plus sensibles au changement et les plus aptes à discriminer dans les domaines prioritaires. Elle avait aussi pour but de mener à un outil qui ne prendrait pas plus de 5 à 10 minutes à compléter. La sélection des questions sensibles au changement et discriminantes a été effectuée en se fondant sur une évaluation approfondie des résultats des tests de sélection, en analysant les différences absolues dans les scores des questions avant et après les traitements et à l'aide du jugement clinique du comité de recherche formé par les co-chercheurs.

Un questionnaire préliminaire de 35 questions a été élaboré. En effet, à cette époque, l'équipe de recherche estimait que le nombre idéal d'items à retenir se situait autour de 35, de façon analogue, par exemple, au SF-36. Certaines questions ont été légèrement reformulées compte tenu que les questions de dépistage pour déterminer l'applicabilité des items n'étaient plus

requises. De plus, il a été décidé de se limiter à deux types d'échelles : une échelle de sept points évaluant le degré de difficulté pour les questions qui portent sur les activités et une échelle de sept points mesurant la fréquence des énoncés pour les autres questions (ex : l'humeur, le rôle social, l'estime de soi, le sommeil). Ainsi, l'échelle pour quelques questions mesurées antérieurement avec une échelle Likert, d'accord et désaccord, a été changée pour une échelle mesurant la fréquence. Cette décision a été prise suite à des données qualitatives provenant des intervieweurs qui ont indiqué que les sujets ayant une faible scolarité éprouvaient de la difficulté à répondre à l'échelle d'accord-désaccord.

2.4 Phase 4 : Traduction en français des items potentiels et des questionnaires de validation

La traduction française des items potentiels, des questionnaires démographiques et de douleur et des questions mesurant l'évaluation globale du sujet et de son thérapeute concernant l'impact de sa lésion sur sa qualité de vie a été effectuée selon une méthode de traduction parallèle à renversement double décrite par Vallerand (1989). Les étapes de cette méthode ont compris :

i. la traduction de la version du questionnaire anglais vers le français par deux traducteurs indépendants (versions A et B) ;
ii. les deux versions françaises (versions A et B) ont été traduites vers l'anglais par deux traducteurs différents (versions C et D) ;
iii. les versions C et D ont été comparées à la version anglaise originale par un comité composé de trois chercheurs bilingues (deux francophones, un anglophone) et les divergences ont été résolues à l'unanimité pour parvenir à une traduction révisée, soit la version E ;
iv. la version E a été prétestée sur 16 travailleurs francophones souffrant de douleurs au cou ou aux membres supérieurs pour identifier les questions ambiguës ou difficiles à comprendre ;
v. les résultats du prétest ont été étudiés par le comité de traduction de recherche et ce dernier s'est entendu sur une version française définitive des questions (version F).

Voici les critères d'acceptation de la formulation française :

- le sens de la version française serait le plus fidèle possible au sens de la version anglaise ;
- les termes les plus simples seraient choisis (pour être compris par un niveau de capacité de lecture de sixième année ou moins) ;
- la syntaxe française serait respectée ;
- le choix se porterait sur les termes les plus connus des québécois francophones.

2.5 Phase 5 : Étude de sélection des items en français

2.5.1 Formation des intervieweurs et tests de fiabilité inter-juges

Préalable à l'étude francophone de sélection des items, des intervieweurs expérimentés ont été recrutés et formés avec les mêmes méthodes de formation que celles utilisées pour l'étude ontarienne. Après la formation, une étude de fiabilité inter-juges a été menée avec 16 patients-travailleurs recrutés dans deux milieux cliniques utilisant les mêmes méthodes que celles décrites précédemment (section 2.2.5).

2.5.2 Population à l'étude

Quatre-vingt-treize (93) travailleurs francophones souffrant de lésions musculo-squelettiques des tissus mous des membres supérieurs ou du cou et répondant aux critères de sélection analogues à l'étude anglophone ont été recrutés dans neuf cliniques privées de physiothérapie et d'ergothérapie de trois régions différentes du Québec (représentant le secteur urbain, de banlieue et semi-rural) de janvier 1996 à juin 1997. Toutes ces cliniques recevaient des patients indemnisés par la Commission de la santé et de la sécurité du travail (CSST) ainsi que d'autres patients. Chaque clinique a suivi ses propres protocoles de traitement. Chaque patient a été vu d'une à trois fois par semaine pour une durée de 20 à 60 minutes. Une séance de traitement comprenait généralement des modalités physiques pour le soulagement de la douleur, l'éducation des patients ainsi que des programmes d'exercice. Comparativement à l'Ontario, une plus grande partie du temps a été accordée aux modalités physiques. Aucun programme de réadaptation de la CSST n'a été strictement défini de façon analogue au programme existant de l'Ontario. Tous les

patients n'ayant pas eu recours à la CSST ont été financés par des régimes d'assurance privée ou ont défrayé les coûts eux-mêmes. Les sujets à l'étude se sont vus rembourser un montant prédéterminé pour leur temps et les frais de stationnement.

2.5.3 Procédures de l'étude

Les procédures d'étude ont été presque identiques à celles de l'étude anglaise à l'exception de ce qui suit. Il a été décidé de limiter le nombre de sujets de l'étude de fiabilité test-retest à 35 sujets. Plutôt que de poser les questions du SIP comme mesure de validation convergente, la mesure générique de l'état de santé SF-36 (Ware et al 1994) a été utilisée. La version préliminaire des 35 questions provenant de l'étude anglophone a été traduite en français et ces questions reformulées ont été posées aux sujets de l'étude québécoise avant l'administration de l'ensemble des 175 items potentiels.

Au Québec, vu la grande distance entre les cliniques, il a été impossible de recourir à des intervieweurs indépendants pour mener les entrevues des questions potentielles et pour administrer les divers tests de validation.

2.5.4 Évaluation et sélection des items de l'outil en français

La même stratégie et les mêmes critères de réduction des items que ceux de l'étude anglophone furent appliqués à l'étude francophone à la fois en ce qui concerne la liste traduite des 175 items potentiels et la liste traduite des 35 items de la version préliminaire du questionnaire décrite à la section 2.3.3.6 et nommé IDVQ-35.

2.6 Phase 6 : Comparaison des résultats des études anglophones et francophones

Avant l'analyse, il a été convenu que les résultats ayant découlé de l'étape de la réduction des questions des études francophones et anglophones seraient comparés et que si les divergences n'étaient pas importantes, un seul questionnaire serait élaboré en combinant les résultats. Par contre, si les résultats de l'étape de la réduction et de la sélection de questions différaient

beaucoup, deux questionnaires seraient élaborés soit un en français reflétant les résultats de l'étude francophone et un en anglais reflétant les résultats de l'étude anglophone.

Les différences des profils démographiques des deux populations étudiées ont d'abord été évaluées selon le sexe, l'âge, le revenu familial, les parties du corps touchées, la durée des symptômes, la situation relative à l'emploi, le statut d'indemnisation par le WCB ou la CSST, et le statut de travail et d'immigration. Les différences ont été évaluées par des «t-tests» ou des analyses de Chi-carré.

Les différences dans les résultats des analyses pour l'élimination et la sélection des questions potentielles entre les deux populations étudiées ont été évaluées en comparant :

1. les scores moyens de classement de chaque domaine entre les deux groupes d'étude à l'aide du test de «Wilcoxon Rank-Sum (Mann-Whitney)» pour les analyses univariées et, pour les analyses multivariées, à l'aide des corrélations partielles de Spearman entre les scores moyens de classement de chaque domaine et le statut du groupe d'étude (c'est-à-dire appartenance au groupe d'étude francophone ou au groupe d'étude anglophone) en contrôlant le sexe, l'âge (deux catégories : moins de 40 ans et de plus de 40 ans), le revenu familial annuel (trois catégories : moins de 20 000 $, de 20 000 à 50 000 $, et plus de 50 000 $) et la durée des symptômes (trois catégories : moins de 3 mois, entre 3 et 6 mois, et plus de 6 mois) ;

2. les scores moyens de la question globale de chaque domaine par analyses de régression multiple dont le score moyen de la question globale de chaque domaine est la variable dépendante, le statut du groupe d'étude (anglais ou français) est la variable indépendante principale et les variables de contrôle incluent le sexe, l'âge, le revenu familial annuel et la durée des symptômes.

Les différences entre les deux populations à l'étude dans les résultats des analyses de réduction des questions pour les questions individuelles ont également été identifiées.

2.7 Phase 7 : Sélection des 20 items appropriés aux deux populations

Les données évaluant les similarités et les différences entre les résultats de l'étude francophone et de l'étude anglophone ont été examinées et interprétées par un sous-comité de co-chercheurs (MJD, PL, JL, SRS) et la question concernant le besoin d'un outil unique comportant les mêmes questions dans les deux langues ou de deux outils distincts comportant différentes questions en français et en anglais a été étudiée et résolue.

Finalement, une sélection définitive de 20 questions pour un seul outil a été effectuée en se fondant sur les résultats obtenus dans les deux populations à l'étude et en s'appuyant sur les critères de sélection mentionnés précédemment (section 2.3.3.6). La formulation des questions de la version IDVQ-35 a été privilégiée pour la version définitive des questions.

L'équipe de recherche a décidé qu'une version plus courte, comprenant 20 questions, plutôt que 35, pourrait s'avérer nettement plus pratique tant dans le contexte clinique qu'en recherche, à la condition que cette version plus courte présente une bonne fiabilité, validité et sensibilité au changement.

Afin de s'assurer que les formulations des questions de la version IDVQ-35 mesuraient les mêmes choses que la version originale des questions, le score moyen des 20 questions des deux versions a été corrélé. De la même façon, les scores moyens des questions définitives sur le travail, des questions sur les activités physiques, des questions psychosociales et des questions sur le sommeil de la version définitive ont été corrélés à leurs versions originales.

Une analyse factorielle des 20 questions reformulées de l'étude québécoise a été effectuée avec rotation «varimax». Les matrices factorielles et les corrélations entre les items et les facteurs ont été évaluées.

2.8 *Phase 8 : Tests de fiabilité, de validité et de sensibilité au changement de la version finale*

L'évaluation de la fiabilité, de la validité et de la sensibilité au changement a été effectuée sur les 20 items sélectionnés. Ces 20 items constituent un sous-ensemble des 175 items potentiels administrés aux sujets francophones et anglophones des deux études. Dans l'étude québécoise, la fiabilité, la validité et la sensibilité au changement de la version reformulée de ces 20 questions ont également été évaluées.

2.8.1 Fiabilité

Deux mesures de fiabilité ont été retenues soit la cohérence interne et la fiabilité test-retest.

2.8.1.1 Cohérence interne

Le coefficient de la cohérence interne indique la corrélation moyenne entre les questions d'un instrument. Elle a été évaluée à l'aide du coefficient Alpha de Cronbach pour 119 sujets anglophones ontariens et 93 sujets francophones québécois présents à l'administration initiale des questionnaires.

2.8.1.2 Fiabilité test-retest

La fiabilité test-retest indique la stabilité de la mesure au cours d'un intervalle assez court pour que le niveau d'incapacité n'ait pas le temps de changer, mais assez long pour que les sujets ne se rappellent pas de leurs réponses antérieures. Elle a été évaluée en calculant, pour chaque question, le coefficient de corrélation intraclasse ($CCI_{1,1}$) (Shrout et Fleiss 1979) entre deux administrations de l'indice, avec un délai de trois à sept jours entre les deux administrations, chez 99 sujets ontariens et 33 sujets québécois.

2.8.2 Validité

2.8.2.1 Validité apparente

La validité apparente traite du point jusqu'auquel l'outil semble mesurer ce qu'il est destiné à mesurer. Afin de s'assurer que les questions de cet outil mesuraient les aspects de la vie vraiment touchés par ce type de lésions musculo-squelettiques, le processus pour identifier les questions potentielles incluait des entrevues auprès de la clientèle cible soit 33 travailleurs-patients ayant des lésions musculo-squelettiques du cou ou des membres supérieurs et un sondage auprès de 30 cliniciens de diverses disciplines en milieu thérapeutique qui soignent des patients ayant ce type de lésions. Ce processus et ses résultats ont déjà été décrits par Stock en 1991.

La validité apparente a été également recherchée dans le processus du choix des questions définitives. Le choix de la version définitive des questions a été influencé par l'importance accordée à chaque domaine par les répondants selon leur classement des domaines et l'impact de leur lésion sur chacun des domaines indiqué par leurs scores moyens de chaque domaine.

2.8.2.2 Validité convergente

La validité convergente démontre le point jusqu'où l'indice est relié aux autres outils qui mesurent la gravité ou le degré d'incapacité associé aux lésions au cou ou aux membres supérieurs. Elle a été évaluée à l'aide des coefficients de corrélations de Pearson entre le score total pour les 20 items sélectionnés pour la version finale de l'outil «IDVQ» et les mesures suivantes : les scores du SIP ou du SF-36, de l'échelle de la douleur, la moyenne des réponses des sujets et des cliniciens à la question mesurant l'évaluation globale du sujet et de son thérapeute concernant l'impact de sa lésion musculo-squelettique sur sa qualité de vie et les résultats aux tests d'examens physiques soit : la force de préhension de la main, l'abduction de l'épaule et le test main-nuque.

2.8.2.3 Validité de construit

Ce type de validité indique le degré avec lequel l'indice fournit une mesure spécifique du construit sous-jacent (état d'incapacité relié aux douleurs au cou ou aux membres supérieurs), relié de façon significative à d'autres mesures de construit théoriquement pertinents. Les chercheurs ont formulé l'hypothèse que les personnes qui étaient au travail lors de la dernière administration des questions «IDVQ», auraient des meilleurs scores IDVQ (i.e. des scores plus bas) que celles qui ne sont pas retournées au travail. Les chercheurs ont également émis l'hypothèse que, parmi les travailleurs qui ont été en arrêt de travail au début des traitements, que ceux qui sont retournés au travail avant la fin des traitements auraient des meilleurs scores IDVQ que ceux qui sont restés en arrêt de travail. Les différences entre les scores ont été vérifiées à l'aide de test (t) de Student.

Dans l'étude québécoise, un questionnaire de détresse psychologique (l'Ilfeld-14) a été administré à tous les répondants. Cet outil a permis la validation de la sous échelle psychosociale de la version définitive de l'IDVQ-20. Les chercheurs avaient formulé l'hypothèse que les six questions de la sous échelle psychosociale de l'IDVQ-20 seraient capables de discriminer entre les répondants démontrant une détresse psychologique élevée au début des traitements selon l'Ilfeld-14 et ceux ne démontrant pas de détresse psychologique. Cette comparaison a été effectuée avec un t test. Le score utilisé comme seuil pour définir une détresse psychologique élevée était fixé au 80^{e} percentile de la distribution des valeurs obtenues à cet indice lors de *l'Enquête Santé Québec 1987* (Boyer et al 1993).

2.8.3 Sensibilité au changement

«La sensibilité au changement» est la capacité d'un outil à détecter correctement une amélioration ou une détérioration de l'état de santé. La sensibilité au changement de la version finale des 20 questions de l'IDVQ a été évaluée auprès de sujets pour lesquels l'évaluation globale du sujet et l'évaluation globale du thérapeute concernant l'impact de la lésion musculo-squelettique sur la qualité de vie indiquaient un changement d'au minimum un point sur l'échelle. Nous avons constaté un changement dans l'état du sujet quand il y avait un consensus dans la présence et la

direction du changement pour ces deux questions. Un tel consensus a été démontré pour 33 sujets ontariens et 35 sujets québécois.

Pour évaluer la sensibilité au changement, des moyennes des réponses normalisées (MRN) ont été calculées pour ces 33 sujets ontariens et 35 sujets québécois. La MRN est une statistique qui mesure la différence entre les scores avant et après les traitements divisée par l'écart type de la différence.

Afin de comparer la sensibilité au changement parmi les différentes mesures de résultats, les moyennes des réponses normalisées ont également été calculées pour les mesures suivantes dans les deux études : le SIP et le SF-36, l'échelle de la douleur, la moyenne des réponses des sujets et des cliniciens à la question mesurant l'évaluation globale du sujet et de son thérapeute concernant l'impact de sa lésion musculo-squelettique sur sa qualité de vie et les tests d'examens physiques soit la force de préhension de la main, l'abduction de l'épaule et le test main-nuque.

Les analyses statistiques ont été effectuées avec les logiciels de statistiques SPSS et SAS.

3. RÉSULTATS

Afin d'en faire une description concise, les résultats des études anglophone (ontarienne) et francophone (québécoise) seront présentés simultanément.

3.1 Identification des items potentiels

L'étude initiale des outils existants, l'examen des documents pertinents et l'expérience clinique des chercheurs principaux ont aidé à identifier 80 questions dans huit domaines, c'est-à-dire dimensions de la vie, où les sujets ont éprouvé de la difficulté en raison de leur problème au cou ou aux membres supérieurs. Le sondage mené auprès des cliniciens a permis d'identifier 52 autres questions ainsi que deux nouveaux domaines, et les entrevues avec les patients-travailleurs ont généré 48 nouvelles questions et deux domaines additionnels. Ce processus a donc mené à l'identification de 12 domaines, soit : le travail, les responsabilités domestiques et familiales, les soins personnels, les activités de loisir, le transport et la conduite de la voiture, la vie sexuelle, le sommeil, la vie sociale, l'humeur, l'estime de soi, les impacts financiers de la lésion et les effets iatrogènes de l'évaluation clinique ou du traitement (Stock et al 1996). Les questions redondantes ont été supprimées et, lorsque jugées appropriées, les questions semblables ont été combinées pour mener à un groupe d'environ 175 items potentiels à inclure dans les études de sélection des questions (Stock 1991).

3.2 Fiabilité inter-juges des items potentiels

3.2.1 Étude anglophone de fiabilité inter-juges

L'étude anglophone de fiabilité inter-juges a révélé que le coefficient de corrélation intraclasse ($CIC_{(1,1)}$) de la moyenne des scores des questions dans chaque domaine était au-dessus de 0,80 dans 11 des 12 domaines. Une CIC moins élevée dans un domaine a été causée par l'interprétation fautive des instructions de codage de l'un des intervieweurs qui a systématiquement mal codé. L'interprétation a été corrigée avant l'étude de sélection des questions.

3.2.2 Étude francophone de fiabilité inter-juges

Dans l'étude de fiabilité inter-juges au Québec, la moyenne des scores des questions de chaque domaine avait un CIC au-dessus de 0,80 dans chacun des 12 domaines. Seulement 13 de l'ensemble des 175 questions individuelles ont eu un CIC en dessous de 0,80. La source de la faiblesse de fiabilité inter-juges pour ces questions individuelles a été identifiée et le problème corrigé. Les sujets qui ont rempli le questionnaire des items potentiels ont très bien accepté le questionnaire et, donc, ont confirmé l'acceptabilité de la durée et du contenu de l'entrevue. Les résultats du CIC de chaque item selon le domaine de l'étude québécoise de fiabilité inter-juges se trouvent à l'Annexe A.

3.3 Sélection des items

3.3.1 Description des populations à l'étude

3.3.1.1 Population de l'étude anglophone

En Ontario, 120 sujets ont été recrutés et ont complété la première série d'entrevues. Au moment de l'analyse, un de ces sujets a été éliminé parce qu'il ne répondait pas au critère de sélection exigeant une compréhension adéquate de l'anglais, laissant 119 sujets pour les analyses des données d'entrevues initiales, et 99 pour les analyses de fiabilité test-retest. Quatre-vingt-sept des 119 sujets (73,1 %) ont complété l'entrevue et les évaluations post-traitement. Parmi les 32 sujets qui n'ont pas complété la dernière entrevue, la raison la plus souvent évoquée consistait à avoir quitté le programme de physiothérapie avant d'avoir complété la série des traitements recommandés ou ne pas avoir eu le temps de participer à l'étude à cause d'engagements au travail ou de responsabilités familiales. Les sujets qui n'ont pas complété la dernière entrevue ne présentaient toutefois pas de différences significatives en fonction du sexe, de l'âge ou de la durée des symptômes comparativement aux 87 sujets ayant complété cette entrevue.

3.3.1.2 Population de l'étude francophone

Au Québec, 93 sujets ont été recrutés et ont complété la première série d'entrevues. 80 des 93 sujets (86%) ont complété l'entrevue et les évaluations post-traitement. Comme prévu au protocole, 35 d'entre eux ont également participé à une entrevue pour l'étude de fiabilité test-

retest (normalement 3 à 7 jours après la première entrevue). Les analyses de deux de ces 35 sujets ont toutefois dû être éliminées en raison d'un délai trop long entre l'entrevue initiale et la seconde entrevue (plus de 7 jours).

Le tableau 1 présente la description sociodémographique des deux populations. La comparaison statistique des deux populations a montré que l'étude francophone a eu plus de sujets féminins (p= 0,02), de sujets ayant des symptômes de plus longue durée (p= 0,001), de sujets au travail (p= 0,02) et moins de sujets recevant des prestations de la CSST ou du « *Workers' Compensation Board* » de l'Ontario (p= 0,0001) que les sujets de l'étude anglophone. La durée moyenne des traitements pour les 87 sujets anglophones qui ont eu une entrevue après les traitements était de 39,7 jours (écart type 14,5 jours; médiane de 41 jours) et était très similaire à la moyenne de 39,0 jours (écart type de 16,1 jours; médiane de 41 jours) des 80 sujets francophones.

Tableau 1 : Comparaison des populations à l'étude

Variables	***Ontario n=119***	***Québec n=93***
Âge moyen (± é.t.)	39,7 ans (± 10,1)	41,1 ans (± 10,0)
	n (%)	**n (%)**
Sexe		
Homme	71 (59,7)	41 (44,1)
Femme	48 (40,3)	52 (55,9)
Durée des symptômes au début de traitement		
< 3 mois	78 (65,5)	31 (33,5)
> 6 mois	34 (28,6)	47 (50,5)
Revenu familial		
< $30,000	44 (36,7)	36 (38,9)
> $50,000	33 (27,5)	34 (36,7)
Statut d'emploi		
Emploi régulier	6 (5,1)	26 (28,0)
Travaux légers	25 (21,2)	14 (15,1)
Arrêt de travail	87 (72,9)	53 (57,0)
Cas indemnisés	81 (67,8)	25 (26,9)
Non syndiqués	67 (56,3)	48 (51,6)
Éducation		
Sans diplôme d'études secondaires	13 (10,9)	18 (19,4)
Diplôme d'études universitaires	4 (3,4)	12 (12,9)

3.3.2 Élimination et sélection des items de l'étude anglophone

Les analyses pour l'élimination des items pour l'étude anglophone ont été effectuées selon les méthodes et les critères décrits au chapitre 2 sur la méthodologie. Une bonne proportion des questions a été éliminée parce qu'elle ne répondait pas au critère d'applicabilité pour les deux sexes, c'est-à-dire que la question n'était pas applicable à au moins 70 % des femmes, 70 % des hommes et 80 % de la population totale à l'étude. Par exemple, une question sur la difficulté de passer l'aspirateur a été éliminée car elle n'était applicable qu'à 49 % des hommes même si elle était applicable à 83 % des femmes. De la même façon, une question sur la difficulté de pelleter de la neige a été éliminée car elle n'était pas applicable à 82 % des femmes. Toutes les questions sur la vie sexuelle ont été exclues car 35 % des femmes ont indiqué qu'elles n'étaient pas actives sexuellement pour des raisons autres que leur lésion musculo-squelettique. Ainsi, cinquante-huit pour-cent (57,6 %) des questions ont été retenues suite à l'évaluation de leur applicabilité.

Les 100 items restants ont été évalués par rapport à leur fiabilité test-retest, leur cohérence interne, leur sensibilité au changement et leur capacité discriminatoire. Seulement 41 items ont répondu à tous ces critères. Parmi ces items, les chercheurs en ont sélectionné une trentaine qui offraient les meilleurs résultats en terme de capacité discriminatoire et de sensibilité au changement. Ces questions ont été adaptées à une version du questionnaire plus raffinée et plus finalisée (voir Annexe B pour la version adaptée de 35 items qui a été évaluée dans l'étude francophone, l'IDVQ-35) c'est-à-dire que les questions ont été reformulées sans les questions d'applicabilité.

Quelques autres constats sont également ressortis de ces analyses :

i. les 3 questions initiales sur la manutention au travail (manutention du plancher à la hauteur de la ceinture, manutention à la hauteur de l'épaule, manutention des personnes) ont indiqué des impacts importants de la lésion sur le travail mais chaque question, prise individuellement, ne répondait pas au critère d'applicabilité chez les femmes. Il a donc été décidé de tester une nouvelle formulation d'une question sur la manutention au travail

qui mesurait le degré de difficulté de «soulever des objets ou des personnes». Cette question a été intégrée au questionnaire IDVQ-35 ;

ii. l'importance de l'impact du problème musculo-squelettique sur les activités de loisir a été cotée assez élevée par les travailleurs masculins et le score moyen du domaine a aussi été élevé (voir Figures 1 et 2 à la section 3.4) mais seule la question globale des activités de loisir répondait au critère d'applicabilité. Les chercheurs ont eu l'impression qu'il y aurait lieu de distinguer parmi les activités de loisir plus actives et les activités plus sédentaires. Il a donc été décidé de séparer ces deux types d'activités de loisir au sein de deux questions et de les intégrer au questionnaire IDVQ-35 afin de les tester dans l'étude francophone. Une question demandait d'indiquer le degré de difficulté éprouvé pour «faire de l'exercice, pratiquer des sports ou d'autres activités récréatives à l'extérieur» et l'autre demandait d'indiquer le degré de difficulté éprouvé pour «faire des activités de loisir, comme de l'art ou de l'artisanat, s'adonner à des hobbys, jouer à des jeux de société ou faire de la musique» ;

iii. Les intervieweurs ont noté, de façon qualitative, que plusieurs répondants de plus faible scolarité éprouvaient de la difficulté à répondre à une échelle de réponse «accord-désaccord» pour les questions psychosociales. Les sujets répondaient avec plus de facilité aux échelles qui mesuraient le degré de difficulté ou la fréquence des items. Selon ce feedback, il a été décidé de tester une version des questions psychosociales en remplaçant l'échelle de réponse «accord-désaccord» par l'échelle de fréquence déjà utilisée pour les questions sur le sommeil, l'impact financier et les effets iatrogènes des évaluations et des traitements. Cette nouvelle version des questions psychosociales a été intégrée à l'IDVQ 35 ;

iv. Pour l'IDVQ-35, il a été nécessaire de reformuler la question globale sur le sommeil et la reformulation suivante a été testée : «je dors mal».

L'annexe C présente les résultats de l'étude anglophone concernant les critères d'élimination et de sélection des items pour les 20 items retenus au cours du processus de sélection définitive ou la version originale de ces questions (Voir section 2.3.3.6 pour l'explication de la différence entre la formulation originale des questions et la formulation dans l'IDVQ-35).

3.3.3 Élimination et sélection des items de l'étude francophone

L'étude francophone a permis l'évaluation de la traduction française des 175 questions potentielles initiales et des 35 questions du format plus raffiné de l'IDVQ-35. Cinquante-trois pour-cent (52,7 %) des 175 questions potentielles, c'est-à-dire 92 items, ont été retenus suite à l'évaluation de leur applicabilité. Par ailleurs, l'ensemble des 35 questions de la version reformulée de l'IDVQ-35 répondait aux critères d'applicabilité.

Dans cette étude, 55 des 175 items répondaient à tous les critères d'élimination et de sélection. Ces 55 items incluaient presque tous les 41 items qui avaient répondu aux critères d'élimination et de sélection de l'étude anglophone. L'Annexe D présente les résultats concernant les critères d'élimination et de sélection des items pour les 20 items retenus au cours du processus de sélection définitive (ou la version originale de ces questions).

Parmi les autres 35 items testés dans la version des questions reformulées de l'IDVQ-35, 27 questions ont répondu à tous les critères (voir Tableau 2). La performance des questions psychosociales avec l'échelle de fréquence a été bonne sur tous les critères à l'exception d'une question qui a eu une fiabilité test-retest faible. Les nouvelles formulations des questions sur la manutention et sur le sommeil mentionnées plus haut ont également affiché de bonnes performances. Mais la nouvelle formulation de la question sur les activités de loisir plus sédentaires s'avérait non applicable pour 27 % des répondants masculins.

Tableau 2 : Résultats de la démarche d'élimination et de sélection des 35 items reformulés de l'IDVQ-35

Item	Applicabilité (% des répondants pour lesquelles l'item a été applicable)		Fiabilité test-retest (Pearson R)	Corrélation entre l'item et le score moyen du domaine sans l'item	Corrélation entre l'item et l'item global du domaine	Skewness (erreur type); Distribution des histogrammes (bien ou mal distribué)	Corrélation entre le CRS de l'item et le CRS de sa sous échelle	Corrélation entre l'item et la question d'évaluation globale par le sujet et le clinicien	Corrélation entre le CRS de l'item et le CRS de la question d'évaluation globale par le sujet et le clinicien
Critères :	>70	>70	>0,5	>0,3	>0,3	Skewness < 2 é-t d'erreur type	>0,3	Corrélation statistiquement significative	>0,3
	Femmes	Hommes	Population totale						
Soins personnels (Q-globale)*	**100**	**100**	**0,59**	**0,72**	**0,68**	**0,38 (,25); bien**	**0,80**	**0,51**	**0,54**
Soins des cheveux	**100**	**100**	**0,79**	**0,72**	**0,66**	**0,18 (,25) bien**	**0,83**	**0,59**	**0,55**
Atteindre le dos	**100**	**100**	**0,76**	**0,77**	**0,71**	**-0,04(25) bien**	**0,80**	**0,58**	**0,63**
Tenir/ utiliser petits objets	100	100	0,75	0,67	0,61	0,63(,25) mal	0,69	0,443	0,37
Se soulever avec les bras	94,2	97,6	0,44	0,69	0,63	-0,05 (,26) bien	0,82	0,47	0,58
Tâches domestiques/ familiales	**98,1**	**100**	**0,54**	**0,67**	**0,60**	**-0,04(,25) bien**	**0,85**	**0,49**	**0,64**
Ouvrir boites de conserve, pots etc.	**100**	**100**	**0,78**	**0,58**	**0,50**	**0,24 (,25) bien**	**0,76**	**0,41**	**0,36**
Porter des sacs, caisses…	**100**	**100**	**0,85**	**0,74**	**0,57**	**-0,13(,25) bien**	**0,88**	**0,53**	**0,61**
Lever les bras pour ranger	100	95,1	0,75	0,72	0,59	-0,5(,25) bien	0,86	0,52	0,55
Faire son travail	100	100	0,81	0,84	0,84	-0,93(,25) mal	0,91	0,61	0,67
Quantité du travail	**96,2**	**100**	**0,86**	**0,85**	**0,83**	**-0,85(,25) acceptable**	**0,92**	**0,69**	**0,70**
Étirer le bras… au travail	**92,3**	**92,7**	**0,82**	**0,82**	**0,73**	**-0,85(,26) acceptable**	**0,94**	**0,53**	**0,74**
Pousser ou tirer des objets au travail	75,0	82,9	0,91	0,78	0,76	-0,98(,28) mal	0,92	0,52	0,69
Plier ou tourner les poignets au travail	76,9	90,2	0,65	0,45	0,29	-0,03 (,27) bien	0,86	0,33	0,72
Soulever des objets ou des personnes	**76,9**	**95,1**	**0,55**	**0,83**	**0,72**	**-0,95(,27) acceptable**	**0,90**	**0,66**	**0,71**
Conduire pendant > ½ hre	92,3	85,4	0,71	0,76	0,71	0,08(,26) bien	0,75	0,59	0,54

Item	Applicabilité (% des répondants pour lesquelles l'item a été applicable)		Fiabilité test-retest (Pearson R)	Corrélation entre l'item et le score moyen du domaine sans l'item	Corrélation entre l'item et l'item global du domaine	Skewness (erreur type); Distribution des histogrammes (bien ou mal distribué)	Corrélation entre le CRS de l'item et le CRS de sa sous échelle	Corrélation entre l'item et la question d'évaluation globale par le sujet et le clinicien	Corrélation entre le CRS de l'item et le CRS de la question d'évaluation globale par le sujet et le clinicien
Faire l'exercice/sports/ loisir à l'extérieur	84,6	95,1	0,71	0,59	0,51	-0,81 (,26) acceptable	0,84	0,48	0,63
Activités de loisir à l'intérieur (artisanat, musique, jeux, hobbys...)	86,5	73,2	0,66	0,54	0,32	0,11(,28) bien	0,82	0,39	0,45
Perte de revenu	80,8	82,9	0,73	0,43	0,72	1,95 (,28) mal	0,42*	0,11 (n.s.)	0,17 (n.s.)
Problèmes d'argent – coûts des soins	86,5	90,2	0,53	0,26	0,46	2,05(,27) mal	0,33*	0,09 (n.s.)	0,20 (n.s.)
Anxieux, nerveux	**100**	**100**	**0,55**	**0,57**	**0,46**	**-0,37 (,25) bien**	**0,87**	**0,26**	**0,64**
Triste, déprimé	**100**	**100**	**0,76**	**0,81**	**0,75**	**0,08(,25) bien**	**0,89**	**0,40**	**0,56**
Irritable, frustré, en colère	**100**	**100**	**0,79**	**0,75**	**0,72**	**0,23(,25) bien**	**0,90**	**0,45**	**0,52**
Disputes avec la famille	100	97,6	0,64	0,58	0,53	0,87(,25) mal	0,74	0,32	0,38
Se fâche rapidement	100	100	0,80	0,66	0,58	0,84(,25) mal	0,84	0,37	0,52
Moins de plaisir	**100**	**100**	**0,77**	**0,61**	**0,52**	**-0,33(,25) bien**	**0,89**	**0,42**	**0,59**
Se sent mal dans sa peau	100	100	0,61	0,58	0,58	-0,13(,25) bien	0,86	0,34	0,53
Se sent moins capable	100	100	0,46	0,61	0,47	-0,39(,25) bien	0,88	0,37	0,65
Ne peut pas prendre soin de la famille	**96,2**	**90,2**	**0,88**	**0,44**	**0,36**	**-0,22(,25) bien**	**0,84**	**0,51**	**0,64**
Peur d'incapacité à gagner sa vie	**100**	**100**	**0,83**	**0,66**	**0,54**	**0,58 (,25) mal**	**0,68**	**0,36**	**0,30**
Difficulté à s'endormir	100	100	0,69	0,63	0,63	-0,24(,25) bien	0,80	0,35	0,58
Se réveiller la nuit	**100**	**100**	**0,79**	**0,73**	**0,73**	**-0,28(,25) bien**	**0,91**	**0,32**	**0,62**
Se réveille de bonne heure	100	100	0,85	0,58	0,58	-0,05(,25) bien	0,89	0,39	0,61

Item	Applicabilité (% des répondants pour lesquelles l'item a été applicable)		Fiabilité test-retest (Pearson R)	Corrélation entre l'item et le score moyen du domaine sans l'item	Corrélation entre l'item et l'item global du domaine	Skewness (erreur type); Distribution des histogrammes (bien ou mal distribué)	Corrélation entre le CRS de l'item et le CRS de sa sous échelle	Corrélation entre l'item et la question d'évaluation globale par le sujet et le clinicien	Corrélation entre le CRS de l'item et le CRS de la question d'évaluation globale par le sujet et le clinicien
Dors mal	100	100	0,86	0,79	0,79	-0,28(,25) bien	0,94	0,44	0,64
Effets négatifs des examens ou traitements	100	97,6	0,76	0,55	0,55	-0,33(,25) bien	0,64*	0,37	0,54

* les items retenus dans la version définitive de l'IDVQ-20 sont indiqués en gras

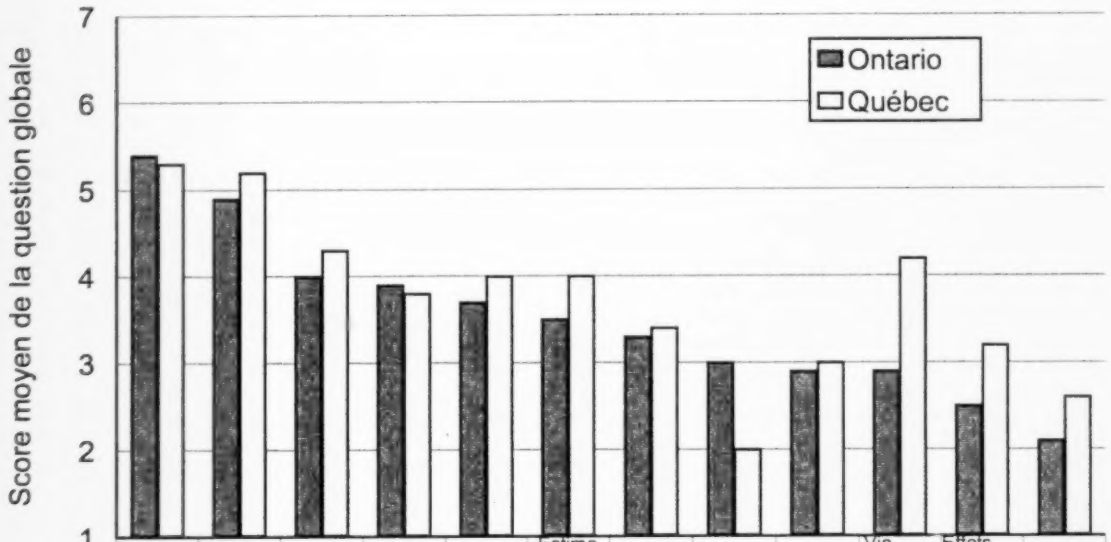
3.4 Importance relative des domaines

La figure 1 présente les scores moyens de la question globale de chaque domaine des sujets anglophones (Ontario) et francophones (Québec). Ces résultats démontrent que, de façon générale, les deux populations à l'étude ont obtenu des scores moyens des domaines très similaires et donc, généralement, les mêmes domaines sont les plus touchés par les LMS du cou et des membres supérieurs. Les domaines avec les scores les plus élevés sont le travail et le sommeil suivis par les activités de loisir, les responsabilités domestiques et familiales et les effets psychosociaux (l'humeur, l'estime de soi). L'impact sur les soins personnels est un peu moins prononcé. Les domaines les moins touchés dans les deux populations sont la vie sociale, la conduite automobile et l'utilisation d'autres moyens de transport de même que les effets financiers des lésions. L'impact sur la vie sexuelle ressort de façon plus importante chez les sujets de l'étude francophone. Dans des analyses selon le sexe, les domaines les plus touchés chez les femmes après l'impact sur le travail et le sommeil sont les responsabilités domestiques et familiales et les effets psychosociaux plus que les loisirs. L'importance de l'impact sur les soins personnels ressort un peu plus chez les personnes qui ont eu des symptômes de longue durée (6 mois et plus).

Les données sur le classement des domaines, présentées à la figure 2*, nous fournissent une autre perspective sur la perception de l'importance des domaines. Généralement les résultats sont très similaires entre les deux populations linguistiques. Le domaine classé le plus important est le travail suivi par le sommeil et les responsabilités domestiques et familiales. Les impacts sur la vie sociale et sur la vie sexuelle sont classés parmi les moins importants dans les deux groupes. Les effets financiers et les effets sur l'estime de soi ont été classés à un rang plus important chez les sujets ontariens que chez les sujets québécois qui ont relégué les effets financiers au dernier rang. Dans les analyses selon le sexe, les écarts de classement les plus importants concernent les activités de loisir (qui ont été classé au 4^{e} rang par les hommes et au 9^{e} rang par les femmes) et l'impact sur les soins personnels classé plus important chez les femmes (4^{e} rang) que chez les hommes (8^{e} rang).

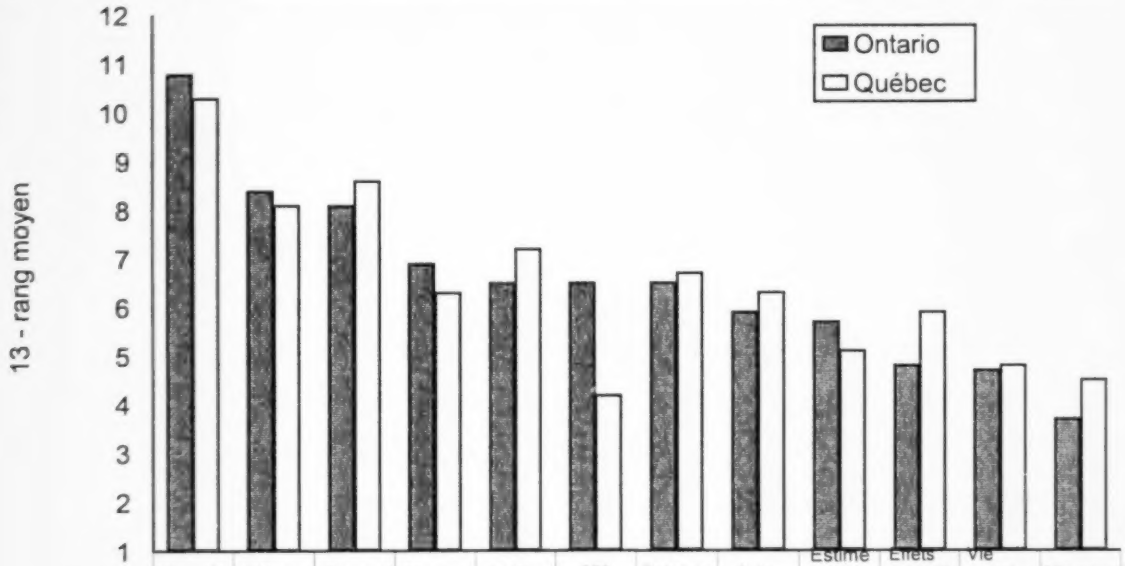
* Le rang moyen le plus petit est celui qui est le plus important. Afin de mieux faire ressortir de façon visuelle l'importance accordée aux rangs des domaines, la figure 2 présente les rangs en hauteur inverse en soustrayant le rang moyen de 13. Ainsi le rang le plus important apparaît le plus haut dans la figure et le rang le moins important apparaît le plus petit.

Figure 1 : Comparaison du score moyen de la question globale de chaque domaine (anglophone vs francophone)



	Travail	Sommeil	Loisir	Humeur	Ménage	Estime de soi	Soins per.	$$$	Auto	Vie sexuelle	Effets négatifs	Vie sociale
Ontario	5.4	4.9	4	3.9	3.7	3.5	3.3	3	2.9	2.9	2.5	2.1
Québec	5.3	5.2	4.3	3.8	4	4	3.4	2	3	4.2	3.2	2.6

Figure 2 : Comparaison du rang moyen de chaque domaine (anglophone vs francophone)



	Travail	Ménage	Sommeil	Humeur	Loisir	$$$	Soins per.	Auto	Estime de soi	Effets négatifs	Vie social e	Vie sexuelle
Ontario	10.8	8.4	8.1	6.9	6.5	6.5	6.5	5.9	5.7	4.8	4.7	3.7
Québec	10.3	8.1	8.6	6.3	7.2	4.2	6.7	6.3	5.1	5.9	4.8	4,5

3.5 Comparaison des résultats des études anglophones et francophones

3.5.1 Comparaison des scores moyens de la question globale de chaque domaine

La figure 1 compare les scores moyens de la question globale de chaque domaine des sujets anglophones (Ontario) et francophones (Québec). Des différences statistiquement significatives ont été observées dans les mêmes trois domaines selon des analyses univariées et multivariées. Le Tableau 3 présente les résultats des régressions multiples qui ont mesuré la relation entre les scores moyens de la question globale de chaque domaine et le statut de groupe d'étude en contrôlant pour le sexe, l'âge, le revenu et la durée des symptômes. Ces résultats montrent que les répondants de l'étude francophone ont obtenu des scores moyens pour la question globale significativement plus élevés pour les domaines de la vie sexuelle ($p = 0,0004$) et des effets iatrogènes des évaluations et des traitements ($p= 0,049$) et que les répondants de l'étude anglophone ont eu des scores moyens pour la question globale du domaine sur l'impact financier de la lésion significativement plus élevés ($p= 0,002$) que les scores des répondants francophones. Pour les autres domaines (neuf) il n'y a pas de différences significatives entre les deux groupes.

Tableau 3 : Relation entre les scores moyens de la question générale de chaque domaine et le statut du groupe d'étude (anglophone vs francophone) en contrôlant pour l'âge, le sexe, le revenu et la durée des symptômes

DOMAINE	Coefficient standardisé	Signif.
Soins personnels	-,016	,829
Responsabilités domestiques et familiales	-,004	,950
Travail	-,041	,587
Transport	-,032	,676
Humeur	-,080	,290
Estime de soi	-,020	,791
Sommeil	-,013	,869
Vie sexuelle	,309[1]	,0004
Loisir	,102	,193
Vie sociale	,108	,156
Effets financiers	-,230[2]	,002
Effets iatrogènes	,149[1]	,049

[1] Un coefficient positif indique que les sujets de l'étude francophone ont eu des scores moyens pour ce domaine significativement plus élevés que les moyens des sujets de l'étude anglophone.
[2] Un coefficient négatif indique que les sujets de l'étude anglophone ont eu des scores moyens pour ce domaine significativement plus élevés que les moyens des sujets de l'étude francophone

Tableau 4 : Comparaison du rang moyen de chaque domaine entre l'étude francophone et l'étude anglophone à l'aide des tests de «Wilcoxon Rank Sum»

DOMAINE	Test du «Wilcoxon rank sum» signif. (p)
Soins personnels	,798
Responsabilités familiales et domestiques	,448
Travail	,099
Transport	,412
Humeur	,168
Estime de soi	,069
Sommeil	,251
Vie sexuelle	,018
Loisir	,084
Vie sociale	,442
Impact financier	,000
Effets iatrogènes	,017

3.5.2 Comparaison du classement des domaines

La figure 2 compare le classement moyen de chaque domaine entre les populations francophones et anglophones. Des différences statiquement significatives de classement des domaines entre les deux groupes à l'étude ont été trouvées lors des analyses univariées, et confirmées par les corrélations partielles de Spearman en contrôlant pour l'âge, le sexe, le revenu et la durée des symptômes. Le Tableau 4 présente les résultats univariés des tests de «Mann-Whitney U-Wilcoxon Rank Sum» qui montrent des différences significatives entre les deux populations dans le classement des mêmes trois domaines soit ceux de la vie sexuelle, des effets financiers et des effets iatrogènes. Dans les analyses multivariées du tableau 5, les sujets de l'étude anglophone ont accordé un classement plus important aux effets financiers reliés aux lésions (p= 0,0001) que les sujets québécois. Par contre, les sujets de l'étude francophone ont accordé un classement plus important aux effets iatrogènes de l'évaluation ou du traitement (p= 0,015), à l'impact des lésions musculo-squelettiques sur les activités de loisir (p= 0,02) et sur la vie sexuelle (p=,051) que les sujets de l'étude anglophone. Aucune différence significative n'a été trouvée dans les autres domaines.

Tableau 5 : Relation entre les scores moyens de classement de chaque domaine et le statut du groupe d'étude (anglophone vs. francophone) en contrôlant pour l'âge, le sexe, le revenu et la durée des symptômes

DOMAINE	Coefficient de corrélation partielle de Spearman	Signif.
Soins personnels	,003	,966
Responsabilités domestiques et familiales	-,037	,622
Travail	-,092	,216
Transport	,076	,313
Humeur	-,116	,122
Estime de soi	-,135	,075
Sommeil	,077	,305
Vie sexuelle	,146	,051
Loisir	,170	,022
Vie sociale	,061	,415
Effets financiers	-,293	,0001
Effets iatrogènes	,180	,015

3.5.3 Interprétation des comparaisons

Suite à la comparaison des résultats obtenus pour chacun des deux groupes culturels, les chercheurs ont convenu que les différences observées n'étaient pas suffisamment importantes pour nécessiter la mise au point de deux outils distincts et ce, pour les raisons suivantes. D'abord, le fait que le groupe ontarien ait accordé une plus grande importance aux répercussions financières était probablement davantage relié au plus grand nombre de sujets en arrêt de travail dans ce groupe qu'à de véritables différences culturelles entre les deux groupes. Ensuite, des variations au niveau des modalités de traitement dispensées à chacun des deux groupes pourraient expliquer les différences observées relativement au classement et au niveau du score moyen attribué aux effets iatrogènes des évaluations médicales et des traitements. Le groupe québécois a probablement reçu des traitements de physiothérapie de nature plus passive et un plus grand nombre d'examens physiques que les sujets ontariens. En effet, le programme ontarien imposé par le *Workers' Compensation Board* de l'Ontario privilégiait surtout l'éducation et les exercices physiques plutôt que les traitements passifs effectués par des physiothérapeutes.

Les sujets de l'étude francophone ont rapporté un score moyen plus élevé que les sujets de l'étude anglophone en ce qui a trait à l'impact de leur problème musculo-squelettique sur la vie sexuelle (4,2 vs 2,9 sur 7) et ont classé ce domaine légèrement plus important (8,5 vs. 9,3 sur 12 domaines). Cette différence pourrait refléter de vraies différences culturelles. Il est possible que les sujets francophones se sentent plus à l'aise de décrire des effets de nature sexuelle que les sujets anglophones. Néanmoins, malgré les différences de classement statistiquement significatives, les sujets francophones ont quand même classé l'impact sur la vie sexuelle au 11ème rang sur 12. De plus, une proportion élevée des deux groupes de sujets ont signalé qu'ils n'étaient pas actifs sexuellement pour des raisons autres que celle liée à leur problème au cou ou aux membres supérieurs si bien que les questions concernant la vie sexuelle n'ont pas satisfait les critères d'applicabilité et n'ont pas été retenues dans la version finale de l'outil. La différence entre les deux groupes quant au classement d'impact de la lésion sur les activités de loisir est statistiquement significative dans les analyses multivariées mais représente une différence relativement mineure (classement moyen de 5,8 vs. 6,5 sur 12 domaines). De plus, il n'y a

aucune différence significative dans le score moyen de la question globale sur les activités de loisir (4,0 vs. 4,3 sur une échelle de sept points).

En outre, aucune différence majeure n'a été identifiée entre les deux groupes quant à la performance des questions individuelles lors de l'analyse d'élimination et de sélection des questions. Il a donc été conclu que : 1) les données n'avaient pas démontré de différences transculturelles importantes quant à l'évaluation de la qualité de vie liée aux lésions musculo-squelettiques au cou ou aux membres supérieurs, et 2) qu'il était possible de choisir des questions communes aux deux groupes de sujets répondant aux critères de sélection.

3.6 Sélection définitive des 20 items appropriés aux deux populations

Selon les résultats sur le classement des domaines et les scores moyens de la question globale de chaque domaine, il a été décidé que l'outil définitif devrait comprendre des items sur le travail, le sommeil, les activités physiques (y compris les responsabilités domestiques et familiales, les soins personnels et le loisir), l'impact psychosocial (y compris les questions sur l'humeur, l'estime de soi et l'exercice des rôles), et les effets iatrogènes de l'évaluation et du traitement. L'évaluation de la performance psychométrique de chacune des questions répondant aux critères de sélection et une discussion avec consensus entre les chercheurs concernant les autres critères mentionnés précédemment ont mené au choix des 20 questions de la version définitive qui se trouve à l'Annexe E (version française) et à l'annexe F (version anglaise). Ces 20 questions comprennent quatre questions sur le travail, sept questions sur les activités physiques, deux sur le sommeil, six sur les effets psychosociaux et une sur les effets iatrogènes de l'évaluation et du traitement.

Dix-neuf des 20 questions de la version définitive sont identiques à la version reformulée des questions de l'IDVQ-35. Dans la version de l'IDVQ-35 deux questions sur les activités de loisirs ont été testées et une de ces questions n'a pas été applicable à 27 % des hommes et 14 % des femmes. Il a été décidé de remplacer ces deux questions par une seule question globale sur les activités de loisir (Voir question 11, Annexe E) qui est très proche de la question globale sur les activités de loisir testée parmi les 175 questions et qui est applicable à presque 100 % des

répondants. Toutes les questions de la version définitive de l'IDVQ-20 sont applicables à plus de 80 % des répondants et 19 parmi les 20 sont applicables à plus de 90 % des répondants.

Afin de s'assurer que les formulations des questions de la version IDVQ-35 mesuraient les mêmes choses que la version originale des questions (voir section 2.3.3.6 pour l'explication de la différence entre la formulation originale des questions et la formulation dans l'IDVQ-35), les scores moyens des 20 questions des deux versions ont été corrélés et ont eu une corrélation de Pearson de 0,98. De la même façon, les scores moyens des quatre questions sur le travail, des sept questions sur les activités physiques, des six questions psychosociales et des deux questions sur le sommeil de la version définitive ont été corrélés à leurs versions originales. Ces analyses ont donné les corrélations suivantes : questions sur le travail - 0,97, questions sur les activités physiques - 0,95, questions psychosociales - 0,96, questions sur le sommeil - 0,91. Ces résultats confirment que les modifications apportées à la version de l'IDVQ-35 n'ont pas changé l'interprétation des questions par les sujets.

3.6.1 Analyses factorielles

Les analyses factorielles avec rotation varimax de la version définitive de l'IDVQ ont confirmé que les questions se distribuent selon quatre facteurs qui correspondent bien aux sous échelles prévues. Les quatre questions sur le travail ont eu des corrélations entre 0,74 et 0,87 avec le premier facteur. Les six questions sur les activités physiques (ce qui n'inclut pas la question sur les activités de loisir) ont eu des corrélations entre 0,47 et 0,78 (0,77, 0,75, 0,78, 0,47, 0,78, 0,50) avec le deuxième facteur ; les 6 questions psychosociales ont eu des corrélations entre 0,32 et 0,87 avec le troisième facteur (0,75, 0,82, 0,74, 0,47, 0,32, 0,53). Les deux questions sur le sommeil ont été corrélées de 0,87 et de 0,82 au quatrième facteur. Ces résultats appuient la présence de quatre sous échelles soit sur les activités physiques, sur le travail, sur les dimensions psychosociales et sur le sommeil.

3.7 Évaluation de la fiabilité de L'IDVQ-20

3.7.1 Cohérence interne

L'IDVQ-20 démontre une forte cohérence interne pour la version originale des questions (Alpha de Cronbach de 0,90 pour les sujets de l'étude anglophone et de 0.92 pour les sujets francophones) ainsi que pour la version reformulée tirée de l'IDVQ-35 (Alpha de Cronbach de 0,93). Voir tableau 6. La version reformulée est celle qui se trouve dans la version définitive de l'IDVQ-20,

Tableau 6 : Fiabilité et sensibilité au changement de l'IDVQ-20

	Ontario (items originaux)	**Québec (items originaux)**	**Québec (items reformulés)**
Reproductibilité fiabilité test-retest ($ICC_{(1)}$)	0,88 (n = 99)	0,85 (n = 35)	0,83 (n = 33)
Cohérence interne (Cronbach alpha)	0,90 (n = 119)	0,92 (n = 93)	0,93 (n = 93)
Sensibilité au changement (Moyenne de réponses normalisées) (MRN) et IC (95 %)	1,48 (1,1 – 1,8) (n = 33)	1,74 (1,4 – 2,1) (n = 35)	1,63 (1,3 – 2,0) (n = 35)

3.7.2 Fiabilité test-retest

Les résultats des analyses de fiabilité pour les échantillons ontariens et québécois sont aussi présentés dans le tableau 6. L'IDVQ démontre des niveaux élevés de fiabilité test-retest pour la version originale des questions (soit une $CIC_{(1,1)}$ de 0,88 pour les sujets de l'étude anglophone et de 0,85 pour les sujets de l'étude francophone) de même que pour la version reformulée dans l'IDVQ-35 administrées auprès des sujets de l'étude francophone (CIC de 0,83).

3.8 Validité de l'IDVQ-20

3.8.1 Validité convergente

Les résultats des analyses pour démonter la validité convergente du score global de l'indice de 20 questions retenues, posées au début des traitements, sont indiqués au tableau 7. On y trouve des corrélations (de Pearson) statistiquement significatives, dans la direction prévue, de toutes les mesures. En générale, les corrélations entre ce score et les autres mesures sont similaires dans les deux études. Néanmoins les 20 questions reformulées (de l'IDVQ-35) posées à l'étude québécoise présentent des corrélations avec les autres mesures un peu plus élevées que la formulation originale des questions dans l'étude ontarienne. Les corrélations entre le score global des 20 questions retenues et les mesures génériques de l'état de santé, soit le SIP et le SF-36, sont de l'ordre de 0,5 à 0,7 ; Les corrélations avec les mesures d'examens physiques du coté atteint se situent entre 0,3 et 0,5 ; les corrélations avec l'échelle de douleur se situent entre 0,4 et 0,6 et celles avec la moyenne du score du sujet et celui de son thérapeute pour la question unique globale sur l'impact du problème musculo-squelettique sur les dimensions importantes de la vie à la maison, au travail ou aux activités de loisir se situent entre 0,6 et 0,73. Il avait été prévu que les corrélations avec les examens physiques oscilleraient autour de 0,3 à cause des différences inter-individuelles importantes inhérentes à ces mesures. Les corrélations obtenues témoignent donc d'une très bonne validité convergente par rapport à ces mesures.

Dans l'étude québécoise, la validité convergente de la sous échelle des questions psychosociales a été évaluée par la corrélation (de Pearson) entre le score de l'échelle et le score sur l'Ilfeld-14 (Boyer et al 1993), un indice de 14 items qui mesure la détresse psychologique. La corrélation a été de 0,75 et, évidemment, hautement significative statistiquement, ce qui suggère que les questions psychosociales de l'IDVQ-20 constituent une bonne mesure de détresse psychologique.

Tableau 7 : Comparaisons des coefficients de corrélation de Pearson entre l'IDVQ-20 et les autres mesures

	Ontario (items originaux)	Québec (items reformulés)
Question globale d'un item (score moyen clinicien-sujet)	0,60	0,73
Échelle de douleur	0,42	0,55
Abduction de l'épaule	-0,32	-0,47
Test « main-nuque »	0,37	0,30
Force de préhension à la main (Jamar)	-0,29	-0,41
SIP	0,66	n/a
SF-36 composant physique	n/a	-0,50
ST-36 composant mental	n/a	-0,52

3.8.2 Validité de construit

La capacité du score moyen des 20 questions de l'IDVQ-20 et de la sous échelle des quatre questions sur le travail à la fin de traitement de discriminer entre les répondants toujours en arrêt de travail à la fin des traitements et ceux retournés au travail a été évaluée. À la fin de l'étude ontarienne, 41 des travailleurs atteints étaient de retour au travail et 44 étaient toujours en arrêt de travail (des données sur le travail manquaient pour 2 travailleurs). À la fin de l'étude québécoise, 51 travailleurs étaient de retour au travail et 29 étaient toujours en arrêt de travail. Les résultats de ces comparaisons statistiques chez ces travailleurs sont présentés au Tableau 8. Les différences statistiquement significatives entre ceux qui étaient au travail et ceux en arrêt de travail démontrent la capacité de discrimination de l'indice et de sa sous échelle du « travail ».

Tableau 8 : Comparaison du score moyen à la fin des traitements, des 20 questions de l'IDVQ-20 et de la sous échelle sur le travail, entre les répondants en arrêt de travail à la fin des traitements et ceux qui étaient au travail.

	Score moyen des répondants au travail	Score moyen des répondants en arrêt de travail	Signif. (p) du t-test*
Étude anglophone			
-Score total IDVQ-20	2,46	3,54	**0,001**
-Sous échelle Travail	3,07	4,16	**0,025**
Étude francophone			
-Score total IDVQ-20	2,20	3,68	**0,000**
-Sous échelle Travail	2,54	5,21	**0,000**

* Les résultats statistiquement significatifs sont en gras

Une autre analyse s'est restreinte aux sujets en arrêt de travail au début des traitements. Elle a permis de tester la capacité de l'IDVQ-20 et de la sous échelle du travail administrée à la fin des traitements de discriminer entre ceux qui sont retournés au travail (en Ontario n=15; au Québec n= 18) et ceux qui sont restés en arrêt de travail avant ce moment (Ontario n=37; Québec n=29). Les résultats statistiques sont présentés au tableau 9. La capacité de l'indice de discriminer entre ces groupes de travailleurs a été validée.

Tableau 9 : Comparaison du score moyen à la fin des traitements, des 20 questions de l'IDVQ-20 et de la sous échelle sur le travail, chez les répondants en arrêt de travail au début des traitements, entre ceux qui sont retournés au travail avant la fin des traitements et ceux qui sont restés en arrêt de travail.

	Score moyen des répondants en arrêt de travail au début des traitements qui sont retournés au travail	Score moyen des répondants en arrêt de travail au début des traitements qui sont restés en arrêt de travail	Signif. (p) du t-test*
Étude anglophone			
-Score total IDVQ-20	2,45	3,54	**0,015**
-Sous échelle Travail	2,52	4,41	**0,001**
Étude francophone			
-Score total IDVQ-20	2,44	3,68	**0,006**
-Sous échelle Travail	2,88	5,21	**0,001**

* Les résultats statistiquement significatifs sont en gras

L'administration d'un indice de détresse psychologique, l'Ilfeld-14, à tous les répondants dans l'étude québécoise, a permis la validation de la sous échelle psychosociale de la version définitive. La comparaison des scores moyens des six questions de la sous échelle psychosociale entre les répondants qui ont démontré une détresse psychologique élevée selon l'Ilfeld-14 au début des traitements (moyenne sur la sous échelle psychosociale de 4,72) et ceux qui n'ont pas démontré une détresse psychologique (moyenne sur la sous échelle psychosociale de 2,98) a révélé une différence statistiquement significative entre les deux groupes ($p<0,0001$). Ce résultat confirme la capacité des questions psychosociales de l'IDVQ-20 de mesurer la détresse psychologique et de discriminer entre les sujets ayant une détresse psychologique élevée et ceux qui ne l'ont pas.

3.9 Sensibilité au changement

Le tableau 6 indique les résultats de l'analyse de la sensibilité au changement de l'IDVQ-20 pour les répondants qui ont démontré un changement d'au moins un point entre le début et la fin des traitements à la question globale unique posée au patient et à son clinicien concernant l'impact de sa lésion sur sa qualité de vie. Les moyennes des réponses normalisées (MRN) sont très élevées et démontrent donc une très forte sensibilité au changement. La MRN pour la version originale des questions administrées aux sujets anglophones est de 1,48 (IC 95 % : 1,1-1,8) et de 1,74 (IC 95 % : 1,4-2,1) pour la même version des questions administrées en français aux sujets de l'étude québécoise. La version reformulée de ces 20 questions (donc version définitive) incluses dans l'IVDQ-35 administrée aux sujets francophones a montré une MRN de 1,63 (IC 95 % : 1,3-2,0). En générale, une MRN de 0,8 ou plus est considérée comme un indicateur de forte sensibilité au changement. Le tableau 10 compare les MRN de toutes les mesures de santé évaluées dans ces deux études. L'IDVQ-20 est la mesure la plus sensible au changement. Néanmoins, l'échelle de douleur, le SIP et la composante physique du SF-36 démontrent une bonne sensibilité au changement dans ces populations atteintes de lésions musculo-squelettiques au cou ou aux membres supérieurs.

Tableau 10 : Comparaison des MRN de l'IDVQ et les autres mesures

	Ontario (items originaux)	**Québec** (items reformulés)
NULI (IDVQ) – 20	1,48 (1,1 – 1,8)	1,63 (1,3 – 2,0)
Échelle de douleur	1,22 (0,9 – 1,6)	1,73 (1,4 – 2,1)
Abduction de l'épaule	-0,61 (-1,0 - -0,3)	-1,16 (-1,6 – 0,7)
Test « main-nuque »	0,02 (0,3 – 0,4)	0,59 (0,2 – 1,0)
Force de préhension à la main (Jamar)	-0,80 (-1,2 - -0,5)	-0,33 (-0,8 - -0,1)
SIP (total)	1,14 (0,8 – 1,5)	n/a
SF-36 composant physique	n/a	-1,26 (-1,6 - -0,9)
SF-36 composant mental	n/a	-0,48 (-0,5 - -0,2)

4. DISCUSSION

Le but général de ce projet était d'élaborer une mesure de l'état de santé fiable et valide, spécifique aux lésions musculo-squelettiques au cou ou aux membres supérieurs, qui pourrait s'appliquer aux populations francophones et anglophones du Canada et s'avérer à la fois utile en recherche et facile d'utilisation pour les milieux cliniques.

L'indice d'impact qui s'est ensuivi, l'IDVQ, est un outil constitué de 20 questions n'exigeant que quelques minutes à remplir, s'appliquant à des populations aux capacités de lecture relativement faibles et offrant la possibilité d'être auto-administré.

Le processus d'élaboration de l'outil, fondé sur les méthodes recommandées par Streiner et Norman (1995) et par Guyatt et al (1986, 1996), a été conçu pour assurer la pertinence clinique maximale et la qualité psychométrique. L'élimination progressive et méthodique des items parmi un grand nombre d'items potentiels a assuré que toutes les questions sélectionnées dans la version définitive répondaient strictement aux critères d'applicabilité, de reproductibilité, de cohérence interne, de sensibilité au changement et de capacité discriminatoire.

Une des forces de l'IDVQ est associée au fait que ses items, générés, entre autres, par des entrevues avec des travailleurs atteints et des sondages auprès des cliniciens qui les traitent, reflètent la réalité vécue par les personnes atteintes. L'importance de tenir compte du vécu des patients dans le développement des outils mesurant l'état de santé et la qualité de vie associés à la santé a été soulignée par Gill et Feinstein (1994) et par Kessler et Mroczek (1995). De plus, les analyses d'applicabilité et de cohérence interne de l'IDVQ ont permis un choix des items qui tient compte de la pertinence et de l'importance de ces items tant pour les travailleuses que pour les travailleurs. Les résultats de ces analyses démontrent également la forte validité apparente de l'outil.

À l'intérieur d'un projet de plusieurs volets, les chercheurs ont mené deux études indépendantes et parallèles pour la sélection des questions et les validations subséquentes dans des populations anglophones et francophones distinctes. Les analyses comparant les échantillons francophones

québécois et anglophones ontariens suggèrent que la base conceptuelle et psychométrique de l'indice est comparable entre ces deux populations. En conséquence, il est fort probable que l'indice aurait une utilité comparable pour des populations anglophones et francophones à travers l'Amérique du Nord. D'autres études devraient vérifier cette hypothèse. L'utilité de cet indice et ses propriétés psychométriques dans des populations de communautés ethniques au Québec, au Canada et aux États-Unis ne peuvent être prises pour acquis et devraient être évaluées (Stock et al 1998).

Les résultats des analyses présentés dans ce document fournissent la preuve que la version définitive de l'IDVQ-20 est, en effet, un indice fiable et valide. La reproductibilité des scores de l'IDVQ est bien démontrée. Son haut niveau de cohérence interne indique que tous les domaines mesurés (dimensions psychologiques et sociales, effets iatrogènes, impacts sur le travail, le sommeil et les activités physiques liées aux soins personnels, aux tâches domestiques et familiales et aux loisirs) sont reliés à l'état de santé général des patients souffrant de LMS du cou et des membres supérieurs. Point encore plus important, il permet d'atteindre des niveaux adéquats de validité et de sensibilité.

Les résultats des analyses de validité indiquent que l'IDVQ est relié de façon significative aux mesures subjectives (question globale unique sur l'impact de la lésion sur les dimensions importantes de la vie de la personne atteinte tel que perçu par le sujet et par son clinicien, échelle de la douleur) et objectives (abduction de l'épaule, test main-nuque, force de préhension de la main). Cet indice est aussi relié aux mesures génériques de l'état de santé utilisées pour chaque population (SIP pour la population anglophone et SF-36 pour la population francophone). L'IDVQ présente donc des avantages comparativement à ces autres instruments, en ce sens qu'il fournit une évaluation complète des impacts reliés spécifiquement aux lésions musculo-squelettiques du cou et des membres supérieurs, dans les domaines de fonctionnement principaux– physiques et psychosociaux, reliés au travail ou non – en un format court et facile à administrer.

Les résultats de la sensibilité au changement ont aussi démontré les avantages de l'IDVQ comparativement aux mesures existantes. Les valeurs de MRN de l'IDVQ ont été excellentes.

Cette mesure a été capable de bien démontrer le changement dans l'état de santé et le fonctionnement avant et après les traitements, pour les sujets ayant eu une amélioration selon un autre indicateur. De plus, l'IDVQ est plus sensible à ces mêmes changements que des mesures objectives d'examens physiques utilisées dans l'étude : abduction de l'épaule, test «main-nuque» et force de préhension de la main. Ces résultats confirment ceux des études précédentes indiquant que les instruments valides et fiables spécifiques à des maladies démontrent une meilleure reproductibilité et sensibilité au changement que les tests cliniques soi-disant « objectifs » comme des tests d'examens physiques, des radiographies ou des tests d'électromyographie (Ahlmen et al 1988, Deyo 1988, Fries 1983, Katz et al. 1994, Levine et al 1993, Wilson et Cleary 1995). Les résultats confirment aussi que les conclusions antérieures voulant que les instruments spécifiques aux lésions musculo-squelettiques soient plus sensibles aux changements que les mesures d'état de santé génériques (Amadeo et al 1996, Bessette et al 1998, Bombardier et al 1995, Patrick et Deyo 1989, Stucki et al 1995). Dans cette étude, le SIP et le SF-36 ont été moins sensibles aux changements que l'IDVQ.

Globalement, ces résultats confirment la fiabilité et la validité de l'IDVQ et démontrent sa capacité à mesurer, dans des études d'intervention, les changements de santé chez chaque sujet au cours du temps et à discriminer entre les sujets.

Les limites principales de ces études sont associées au fait que la sélection des items a été réalisée dans les mêmes études et les mêmes populations que l'évaluation de la fiabilité et la validation. Afin d'évaluer l'applicabilité des items sélectionnés, il a été nécessaire de tester une formulation des items qui a été légèrement différente de la formulation définitive élaborée par les chercheurs de l'IDVQ. Heureusement, il a été possible également de tester la formulation révisée dans l'étude québécoise. Les corrélations des scores des deux versions ont démontré que les deux versions des 20 questions sélectionnées pour la version définitive sont presque identiques (ex. : corrélation du score total de 0,98). De la même façon, les résultats des analyses concernant la cohérence interne, la fiabilité test-retest, la sensibilité au changement et la validité convergente de la version des questions reformulées sont très similaires aux analyses des items originaux. Ceci suggère que la version définitive recommandée par les auteurs présente une forte fiabilité et sensibilité au changement.

Il existe un certain nombre de limites inhérentes à l'utilisation de l'IDVQ. La formulation des questions de l'IDVQ demande au répondant de juger si ses difficultés aux diverses activités ou ses émotions négatives résultent de sa lésion musculo-squelettique. Ceci est nécessaire pour mesurer l'impact de la lésion sur ces dimensions de la vie. Mais il est possible que certaines personnes éprouvent des difficultés à évaluer l'attribution causale de leur problème de santé. Aussi, comme toute mesure subjective, il est possible que les perceptions de la personne atteinte soient influencées par des caractéristiques personnelles qui ne sont pas nécessairement liées à la santé (ex : personnalité, motivation, humeur). Néanmoins, plusieurs études ont démontré la validité et la valeur prédictive des mesures subjectives de l'état de la santé. Comme indiqué plus haut, elles sont souvent plus fiables et plus sensibles au changement que les mesures plus objectives.

Quand ce projet a débuté, aucune mesure d'état de santé spécifique aux LMS du cou et des membres supérieurs n'existait. Depuis ce temps, au moins deux outils ont été élaborés et validés par d'autres groupes de chercheurs. L'outil le plus développé et validé est le «*DASH : Disability of the Arms, Shoulders and Hands*» conçu par un groupe de chercheurs de Toronto en collaboration avec l'American Academy of Orthopedic Surgeons (McConnell et al 1999, Davis et al 1999, Beaton et al 2001). Cet outil est similaire à l'IDVQ mais avec quelques différences importantes. Il comporte 30 items de base avec quatre items facultatifs pour le travail et est donc plus long que l'IDVQ sans offrir de meilleures caractéristiques psychométriques. Il n'inclut presque aucun item sur les dimensions psychosociales – il a été conçu pour l'administration avec le SF-36 afin de mesurer les aspects psychologiques associés aux LMS. Il se limite aux membres supérieurs et n'a pas été conçu pour les lésions du cou. Plusieurs travailleurs souffrent de lésions brachiocervicales qui touchent la région cervicale et l'épaule. Il s'avère donc intéressant de disposer d'un outil qui s'applique, en même temps, aux personnes atteintes au cou et aux membres supérieurs. Aussi, le processus d'élaboration du DASH n'a pas évalué l'applicabilité des items et les différences selon le sexe de l'applicabilité ou de l'importance accordée aux items. Il a été élaboré dans des populations anglophones ontariennes et américaines et non pas dans des populations francophones. Une traduction validée en français et une étude de validation est en cours au Québec. La chercheure principale de l'IDVQ est à mener une étude qui compare

la sensibilité au changement et la validité de convergence de l'IDVQ et du DASH dans des populations cliniques atteintes de traumatismes de la main et du poignet.

L'autre outil mesurant le statut fonctionnel des personnes atteintes aux membres supérieurs décrit dans la littérature scientifique est celui de Pransky et al (1997). Cet outil de huit items d'activités de vie quotidienne touche très peu de domaines. Il a montré une bonne capacité discriminatoire, une validité convergente et une sensibilité au changement chez des travailleurs américains atteints du syndrome du canal carpien et d'autres LMS du membre supérieur.

Le seul outil existant qui a l'air d'être similaire à l'IDVQ est le DASH. Il serait intéressant de comparer les caractéristiques psychométriques des deux outils dans des populations différentes afin d'avoir une meilleure idée à savoir quel outil sera le plus approprié pour différentes études et contextes cliniques.

5. CONCLUSION

L'IDVQ-20 a été développé selon des méthodes scientifiques rigoureuses et à point. La version définitive comprend 20 questions évaluant les impacts des lésions musculo-squelettiques du cou et des membres supérieurs sur le travail, les activités physiques, le sommeil, les dimensions psychosociales ainsi que les effets iatrogènes chez les personnes atteintes.

Les résultats des études de validation ont démontré que l'IDVQ-20 possède une forte fiabilité test-retest ($CIC_{(1,1)}$ de 0,88 à l'étude anglophone et 0,83-0,85 à l'étude francophone) et cohérence interne (Alpha de Cronbach de 0,90 pour les sujets de l'étude anglophone et de 0,92 à l'étude francophone) aussi bien qu'une sensibilité au changement très robuste (moyenne des réponses normalisées de 1,48 pour la version anglaise et de 1,63 pour la version française). Cet outil présente également une bonne validité convergente selon les corrélations à plusieurs mesures objectives et subjectives d'examens physiques, de douleur et de mesures génériques de l'état de santé. Dans des études de validation de construit, il a aussi été capable de discriminer de façon significative entre les personnes absentes du travail et celles qui travaillaient. La sous échelle des questions psychosociales démontre une bonne capacité de discriminer entre les sujets ayant une détresse psychologique élevée et les autres sujets.

L'IDVQ-20 se présente sous un format court et facile à compléter. Il fournit une évaluation fiable et valide des impacts spécifiquement liés aux lésions musculo-squelettiques du cou et des membres supérieurs dans les principales dimensions de la vie touchées par ces lésions chez les travailleurs atteints de ces lésions des tissus mous.

6. APPLICABILITÉ DES RÉSULTATS ET RETOMBÉES ÉVENTUELLES

L'IDVQ peut servir à plusieurs applications. Il permet d'évaluer rapidement la gravité des incapacités et d'un handicap liés aux lésions musculo-squelettiques d'origine professionnelle.

Dans le contexte de la recherche, cet outil a démontré toutes les caractéristiques psychométriques et «clinimétriques» pour être considéré comme une excellente mesure de résultat dans des études d'évaluation, d'efficacité, d'intervention ainsi que les études de prévalence, de pronostic et d'identification des facteurs de risque des LMS du cou ou des membres supérieurs.

Dans le contexte de l'intervention clinique en médecine de réadaptation, d'orthopédie, de rhumatologie et de médecine du travail ainsi qu'en physiothérapie et en ergothérapie, cet outil permettra un suivi clinique rapide et valide de l'évolution du patient. Avant d'intervenir, cet outil pourrait aider dans le choix du type de mesures de réadaptation. Les sous échelles de l'outil peuvent servir comme outil de dépistage afin d'identifier les patients qui nécessitent des interventions concernant le travail ou le sommeil ou qui ont besoin d'un support psychologique.

L'outil IDVQ peut également servir à des intervenants et des chercheurs de santé publique dans des études de surveillance active afin d'évaluer la gravité des LMS identifiées par des questionnaires de symptômes. Des ergonomes ou d'autres intervenants en santé au travail peuvent aussi le trouver utile et pertinent pour évaluer l'évolution de la gravité des lésions au cou ou aux membres supérieurs dans des études d'interventions ergonomiques ou pour documenter les impacts des lésions.

7. LISTE DES ARTICLES SCIENTIFIQUES PRODUITS DANS LE CADRE DE CETTE SUBVENTION

Stock SR, Loisel P, Durand MJ, Streiner D, Tugwell P, Lemaire J, Boucher M. Cross-cultural Differences Between English and French Canadians and the Measurement of Health Status of Workers with Neck and Upper Limb Disorders. PREMUS-ICEOH '98 Conference, Helsinki, Finland, September 1998 Book of Abstracts, p. 20.

Stock SR, Loisel P, Durand MJ, Streiner D, Tugwell P, Lemaire J, Boucher M. Cross-cultural differences between English Ontarians and French Québécois: Do they influence measurement of functional status of workers with neck and upper limb disorders? *Canadian Journal of Rehabilitation* 1998, Vol. 11, 270-271.

Stock SR, Katz J, Beaton D. Measuring severity and disability in neck and upper extremity disorders: a critical review of existing instruments, Document prepared for Consensus Conference on Case Definitions of Upper Extremity Musculoskeletal Disorders, Nov. 15-16, 1996, Center for VDT and Health Research, Johns Hopkins University, Baltimore, MD, 1-30.

Stock SR, Cole DC, Tugwell P. Review of applicability of existing functional status measures to the study of workers with musculoskeletal disorders of the neck and upper limb. *American Journal of Industrial Medicine* 1996, 29, 679-688

Stock SR, Streiner D, Tugwell P, Loisel P, Reardon R, Durand M-J, Abrahamowicz M, Lemaire J. Validation of the neck and upper limb index (NULI), a functional status instrument for work-related musculoskeletal disorders. Book of Abstracts: ICOH (International Commission on Occupational Health) '96 Conference, Stockholm, Sweden, Sept. 1996

Stock SR, Streiner D, Reardon R, Darzins S, Dilworth P, Tugwell P, Loisel P. The impact of neck and upper limb musculoskeletal disorders on the lives of affected workers: development of a new functional status index. *Quality of Life Research* 1995, 4, 491

Stock SR, Streiner D, et al. The impact of neck and upper limb disorders on the lives of affected workers: development of a new functional status index, Book of Abstracts PREMUS 95, Montreal, Sept. 1995, p. 99-101

Stock SR. Measuring functional status in workers with neck and upper limb disorders: (1) Review of existing instruments; (2) Development of a new instrument. Book of Abstracts: Outcome Research in Work-related Musculoskeletal Disorders. (Sponsor: University of Massachusetts at Worcester and NIOSH). Denver, Co, Oct. 1994

En préparation :

Stock, S, Loisel, P, Durand, M-J, Streiner, D, Tugwell, P, Reardon R, Lemaire, J, Abrahamowicz M, Boucher, M, Darzins S, Dilworth P, Goudreau N. *The Neck and Upper Limb Index* (NULI):

1) Part 1.: Development of a New Musculoskeletal Disorders-Specific Health Status Instrument; The Neck and Upper Limb Index (NULI):

2) Part 2.: Reliability, Sensitivity to Change and other Validity Testing

RÉFÉRENCES

Ahlmen M, Sullivan M, Bjelle A. Team versus non-team outpatient care in rheumatoid arthritis: a comprehensive outcome evaluation including an overall health measure. *Arthritis and Rheumatism* 1988;140: 471-9

Amadeo PC, Silverstein MD, Ilsttrup DM, Schleck CD, Jensen LM. Outcome measures for carpal tunnel surgery: the relative responsiveness of generic, arthritis-specific, disease specific and physical examination measures. *Journal Hand Surgery* 1996; 21A: 338-46.

Arcand R, Labrèche F, Stock S, Messing K, Tissot F. Chapitre 26: Travail et santé. Institut de la statistique du Québec (ed). L'Enquête sociale et de santé 1998, Les publications du Québec, Ste-Foy, Québec, 2000

Beaton DE, Katz JN, Fossel AH, Wright JG, Tarasuk V, Bombardier C. Measuring the whole or the parts? Validity, reliability, and responsiveness of the Disabilities of the Arm, Shoulder and Hand outcome measure in different regions of the upper extremity. *Journal Hand Therapy* 2001 Apr-Jun;14(2):128-46

Bell MJ, Bombardier C, Tugwell P. Measurement of functional status, quality of life and utility in rheumatoid arthritis. *Arthritis and Rheumatism* 1990; 33: 591-601.

Bergner M, Bobbitt RA, Carter WB, Gilson BS. The Sickness Impact Profile: development and final version of a health status measure. *Medical Care* 1981; 19: 787-805.

Bergner M. Quality of life, health status and clinical research. *Medical Care* 1989; 27: S148-56

Bessette L, Sangha O, Kuntz KM, Keller RB, Lew RA, Fossel AH, Katz JN. Comparative responsiveness of generic versus disease-specific and weighted versus unweighted health status measures in carpal tunnel syndrome. *Medical Care* 1998; 36: 491-502.

Bombardier C, Melfi CA, Paul J, Green R, Hawke G, Wright J, Coyte P. Comparison of generic and disease–specific measure of pain and physical function after knee replacement surgery. *Medical Care* 1995; 33: AS131-44

Boyer R, Préville M, Legaré G, Valois P. La détresse psychologique dans la population du Québec non institutionnalisée : résultats normatifs de l'enquête Santé Québec. *Revue canadienne de psychiatrie* 1993; 38: 339-343

Buckle P, Devereux J. Work-related Neck and Upper Limb Musculoskeletal Disorders. Agence européenne pour la sécurité et la santé au travail, Luxembourg, 1999, 117 pages

Cohen ME et Marino RJ. The tools of disability outcomes research functional status measures. *Archives of Physical Medicine and Rehabilitation* 2000; 81 (suppl 2): S21-S29

Davis AM, Beaton DE, Hudak P et al. Measuring disability of the upper extremity: a rationale supporting the use of a regional outcome measure. *Journal Hand Therapy* 1999; 12: 269-274

Deyo R, Carter WB. Strategies for improving and expanding the application of health status measures in clinical settings: a researcher-developer viewpoint. *Medical Care* 1992; 30(suppl) MS176-MS186

Deyo RA. Measuring the functional status of patients with low back pain. *Archives of Physical Medicine and Rehabilitation* 1988; 69: 1044-53

Feinstein AR, Josephy BR, Wells CK. Scientific and clinical problems in indexes of functional disability. *Annals of Internal Medicine* 1986; 105: 413-20

Francis DJ, Fletcher JM, Stuebing KK et al. Analysis of change: modeling individual growth. *Journal of Consulting Clinical Psychology* 1991; 59: 27-37

Fries JF. Toward understanding of patient outcome measurement. *Arthritis and Rheumatism* 1983; 26: 697-704

Gill TM, Feinstein AR. A critical appraisal of the Quality-of-Life measurements. *Journal of American Medical Association* 1994; 272: 619-626.

Greenfield S, Nelson EC. Recent developments and future issues in the use of health status measures in clinical settings. *Medical Care* 1992; 30(suppl): MS23-MS41

Guillemin F. Functional disability and quality-of-life assessment in clinical practice. *Rheumatology* 2000; 39: 17-23

Guyatt GH, Bombardier C, Tugwell PX. Measuring disease-specific quality of life in clinical trials. *Canadian Medical Association Journal* 1986; 134: 889-895

Guyatt GH, Jaeschke R, Feeny DH, Patrick DL. Measurement in clinical trials: Choosing the right approach. In Spilker B (ed), Quality of Life and Pharmacoeconomics in Clinical Trials, Second edition. Lippencott-Raven Publishers, Philadelphia, 1996.

Jette AM. Concepts of health and methodological issues in functional assessment. In Granger CV, Gresham GE, ed. Functional Assessment in Rehabilitation. Williams and Wilkins, Baltimore, 1984; 46-63.

Juniper EF, Guyatt GH, Jaeschke R. How to develop and validate a new health-related quality of life instrument. In Spilker B (ed), Quality of Life and Pharmacoeconomics in Clinical Trials, Second edition. Lippencott-Raven Publishers, Philadelphia, 1996.

Kaegi E. New approaches to Worker Compensation. *Occupational Health in Ontario* 1990; 11: 61-70

Katz, JN, Gelberman, RH, Wright, EA, Lew, RA, Liang, MH. Responsiveness of self-reported and objective measures of disease severity in carpal tunnel syndrome. *Medical Care* 1994; 32 (11): 1127-33.

Kessler RC, Mroczek DK. Measuring the effects of medical interventions. *Medical Care* 1995; 33: AS109-119

Kirshner B, Guyatt G. A methodological framework for assessing health indices. *Journal of Chronic Diseases* 1985; 38: 27-36

Levine, DW, Simmons, BP, Koris, MJ, Daltroy, LH, Hohl, GG, Fossel, AH, Katz, JN. A self-administered questionnaire for the assessment of severity of symptoms and functional status in carpal tunnel syndrome. *Journal of Bone and Joint Surgery* 1993; 75A (11): 1585-92.

McConnell S, Beaton D, Bombardier C. The DASH Outcome Measure User's Manual. Institute for Work and Health, Toronto, and American Academy of Orthopedic Surgeons, Rosemont, IL, 1999

Mitchell RI, Carmen GM. Results of a multicenter trial using an intensive active exercise program for the treatment of acute soft tissue and back injuries. *Spine* 1990, 15(6): 514-21.

Muldoon MF, Barger SD, Flory JD, Manuck SB. What are quality of life measurements measuring? *British Medical Journal* 1998; 316: 542-5.

National Research Council. Panel on Musculoskeletal Disorders and the Workplace, Commission on Behavioral and Social Sciences and Education. Musculoskeletal Disorders and the Workplace: Low Back and Upper Extremities. National Academy Press, Washington DC, 2001, 450 pages

Naughton MJ, Shumaker SA. Assessment of health-related quality of life in orthopaedic outcome studies. Arthroscopy: *Journal Arthroscopic and Related Surgery* 1997; 13: 107-113

Nishisato N, Torii Y. Effects of categorizing continuous normal distributions on the product-moment correlation. *Japanese Psychological Research* 1970, 13: 45-9

Norman GR. Issues in the use of change scores in randomized trials. *Clinical Epidemiology* 1989; 42(11): 1097-105

Nunnally JC. Psychometric Theory, Second Edition. 1978, McGraw Hill, New York, 66-67

Nunnally JC. The study of change in evaluation research: principles concerning measurement, experimental design and analysis. In Struening EL and Guttentag M (ed), Handbook of Evaluation Research, Vol.1. Sage publications, Beverley Hills, CA, 1975

Nunally JC, Wilson WH. Method and theory for developing measures in evaluation research. In Struening EL and Guttentag M (ed), Handbook of Evaluation Research, Vol.1. Sage publications, Beverley Hills, CA, 1975

Patrick DL. Finding health-related quality of life outcomes sensitive to health care organisation and delivery. *Medical Care* 1997; 35: NS49-57

Patrick DL, Deyo RA. Generic and disease-specific measures in assessing health status and quality of life. *Medical Care* 1989; 27 (suppl): S217-232

Pransky G, Feuerstein M, Himmelstein J, Katz JN, Vickers-Lahti M. Measuring functional outcomes in work-related upper extremity disorders. *Journal of Occupational and Environmental Medicine* 1997; 39: 1195-1202

Shrout PE, Fleiss JL. Intraclass correlations: assessing rater reliability. *Psychological Bulletin* 1979; 86: 420-8.

Stock SR, Cole DC, Tugwell P. Review of applicability of existing functional status measures to the study of workers with musculoskeletal disorders of the neck and upper limb. *American Journal Industrial Medicine* 1996, 29, 679-688

Stock SR, Loisel P, Durand MJ, Streiner D, Tugwell P, Lemaire J, Boucher M. Cross-cultural differences between English Ontarians and French Québécois: Do they influence measurement of functional status of workers with neck and upper limb disorders? *Canadian Journal of Rehabilitation* 1998, Vol. 11, 270-271

Stock SR. Measurement of Functional Status in Workers with Musculoskeletal Disorders of the Neck and Upper Limb, M.Sc. Thesis, McMaster University , Hamilton, Ontario, 1991

Streiner DR. Chapter 7: Global Rating Scales, in Neufeld VR, Norman GR (ed): Assessing Clinical Competence. Springer Publishing Co, New York, 1985

Streiner DL, Norman GR. Health Measurement Scales: A Practical Guide to Their Development and Use. Second Edition. Oxford Medical Publications, London, 1995

Stucki G, Liang MH, Fossel AH, Katz JN. Relative responsiveness of condition-specific and generic health status measures in degenerative lumbar stenosis. *Journal Clinical Epidemiology* 1995; 48: 1369-78

Testa MA, Simonson DC. Assessment of Quality of Life outcomes. *New England Journal Medicine* 1996; 334: 835-840

Tugwell P, Bombardier C. A methodologic framework for developing and selecting endpoints in clinical trials. *Journal of Rheumatology* 1982; 9: 758-761

Vallerand, R Vers une méthodologie de validation trans-culturelle de questionnaires psychologiques: implications pour la recherché en langue française. *Canadian Psychology* 1989; 30(4): 662-80.

Van Schoor JT, Wright GR. Rehabilitation of the injured worker. *Canadian Family Physician* 1989; 35: 2297-2300

Ware JE. Conceptualizing disease impact and treatment outcomes. *Cancer* 1984; 53 (Suppl): 2316-26.

Ware JE, Kosinski M, Keller SD. SF-36 Physical and Mental Health Summary scales: A User's Manual. The Health Institute, Boston, MA, 1994.

Ware JE. The status of health assessment 1994. *Annual Review of Public Health* 1995; 16: 327-354

Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. *Journal of American Medical Association* 1995; 273(1): 59-65.

Annexe A

Les résultats de fiabilité inter-juges (ICC) de chaque item selon le domaine

Soins personnels

1	.94
2	.82
3	.98
4	.96
5	.97
6	.85
7	.91
8	.89
9	.96
Question globale	.95

La vie sexuelle

1	.87
2	.91
3	.97
4	1.0
5	.60
6	.91
7	.95
Question globale	.88

Responsabilités domestiques et familiales

1	.92
2	.95
3	.98
4	.93
5	.98
6	.96
7	.89
8	.98
9	.84
10	.87
11	.89
12	.94
13	.99
14	.99
15	.78
16	.90
17	1.0
18	.98
19	.99
20	.71
21	.50
22	.50
Question globale	1.0

Travail

1	.94
2	.95
3-0	.91
3-1	.98
3-2	.95
3-3	1.0
4	.98
5	.97
6	.93
7	.89
8	.95
9	.78
10	.98
11	.99
12	.94
13	.95
14	.99
15	1.0
16	.99
17	1.0
18	1.0
19	.79
20	.85
21	.98
22	1.0
23	1.0
24	.85
Question globale	.98

Activités sociales

1	1.0
2	.95
3	1.0
4	.99
5	.95
6	.97
Question globale	.97

Sommeil

1	.92
2	.99
3	.98
4	.95
5	1.0
Question globale	.99

Humeurs

1	.98
2	.99
3	.82
4	.95
5	1.0
6	.89
7	1.0
8	.60
8.1	.92
9	.82
10	.96
11	.92
12	1.0
13	.92
14	.88
15	.99
16	.98
Question globale	.90

Estime de soi

1	.89
2	.99
3	.92
4	.99
5	.76
6	.96
7	.98
8	.86
9	.99
10	.99
11	.99
12	.78
13	.98
14	.94
15	.99
16	.99
Question globale	.88

Activités de loisirs

1	1.0
2	1.0
3	.98
4	.66
5	.81
6	.97
7	1.0
8	.70
9	1.0
10	1.0
11	.99
12	.99
13	1.0
14	.99
15	1.0
16	1.0
17	1.0
18	1.0
19	1.0
20	.99
21	.90
22	.93
23	.57
24	.91
25	1.0
26	1.0
Question globale	.94

Conduite automobile et autres moyens de transport

A	1.0
1	.93
2	.98
3	1.0
4	1.0
5	.99
6	.97
7	1.0
8	1.0
9	.99
10	.96
11	.97
12	.99
13	.90
14	1.0
15	.92
Question globale	.97

Effets de la lésion sur la situation financière

1	1.0
2	.89
3	1.0
4	.93
5	1.0
6	.97
8	N/A
9	1.0
9.1	1.0
10	1.0
11	1.0
Question globale	.95

Effets secondaires négatifs des examens et traitements

1	1.0
2	1.0
3	.99
4	.98
5	.97
6	1.0
Question globale	.95

Annexe B

IDVQ-35

Indice d'Impact des Douleurs du cou et des membres supérieurs sur la Vie Quotidienne

PREVICAP
Centre de recherche
Hôpital Charles-LeMoyne

DE MONTRÉAL-CENTRE
Direction de la santé publique

INDICE D'IMPACT DES DOULEURS DU COU ET DES MEMBRES SUPÉRIEURS SUR LA VIE QUOTIDIENNE (IDVQ-35)

Les réponses à ce questionnaire sont confidentielles.
Aucune information susceptible de vous identifier ne sera dévoilée.

INSTRUCTIONS

Les questions suivantes nous aideront à mieux comprendre votre problème (douleurs, courbatures, inconforts) du cou ou des membres supérieurs (ex: épaules, coudes, poignets, mains).

Pour chaque question veuillez cocher une réponse de "**1**" (vous n'avez aucune difficulté) à "**7**" (vous ne pouvez pas le faire) selon l'importance que vous jugez. Prenez note que pour la Section B (questions 21 à 35), les chiffres de "1" à "7" mesurent la fréquence (ex: "1" = jamais et "7" = tout le temps).

Il faut répondre "ne s'applique pas" si:

- une question ne vous concerne pas ou
- vous n'êtes pas certain(e) de la réponse ou
- la situation décrite dans la question n'est pas liée à votre problème.

SECTION A

Quel degré de difficulté avez-vous ou auriez-vous *MAINTENANT* avec les activités ou les situations suivantes, À CAUSE DE VOTRE PROBLÈME?
MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS
Aucune difficulté	Un peu difficile	Assez difficile	Moyennement difficile	Très difficile	Extrêmement difficile	Je ne peux pas le faire, ou c'est un problème très grave	Cochez "ne s'applique pas" si la question ne vous concerne pas ou si vous ne savez pas

1 Vous laver, vous habiller, vous arranger ou manger?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

2 Vous laver le dessus de la tête ou laver, sécher ou arranger vos cheveux?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

3 Atteindre votre dos (par exemple pour laver ou essuyer votre dos, pour sortir votre portefeuille de votre poche arrière ou pour attacher votre soutien-gorge ou une fermeture éclair)?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

4 Tenir et utiliser de petits objets (comme tenir une tasse de café sans en renverser, couper de la viande ou d'autres aliments avec un couteau et une fourchette, vous maquiller)?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

5 Vous soulever avec vos bras (par exemple pour sortir du lit, pour vous lever d'une chaise ou d'un fauteuil bas ou pour sortir du bain)?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

6 Faire les tâches à la maison ou vous occuper de votre famille?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

7 Ouvrir des boîtes de conserve ou des pots, tourner des clés, des robinets ou des poignées de porte?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

Quel degré de difficulté avez-vous ou auriez-vous *MAINTENANT* avec les activités ou les situations suivantes, À CAUSE DE VOTRE PROBLÈME?
MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS
Aucune difficulté	Un peu difficile	Assez difficile	Moyennement difficile	Très difficile	Extrêmement difficile	Je ne peux pas le faire, ou c'est un problème très grave	Cochez "ne s'applique pas" si la question ne vous concerne pas ou si vous ne savez pas

8 Transporter des sacs d'épicerie, des caisses de bière ou de boissons gazeuses, ou sortir les déchets de la maison?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

9 Lever les bras pour ranger l'épicerie ou de la vaisselle dans les armoires de la maison?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

10 Faire votre travail?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

11 Faire la *quantité* de travail que l'on attend de vous?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

12 À votre travail, étirer les bras vers l'avant, vers l'arrière, vers le haut, vers le bas ou sur les côtés, de façon répétitive ou en forçant?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

13 À votre travail, pousser ou tirer des objets?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

14 À votre travail, plier ou tourner vos poignets de façon répétitive?

1 2 3 4 5 6 7 — NE S'APPLIQUE PAS

Quel degré de difficulté avez-vous ou auriez-vous *MAINTENANT* avec les activités ou les situations suivantes, À CAUSE DE VOTRE PROBLÈME?
MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS
Aucune difficulté	Un peu difficile	Assez difficile	Moyennement difficile	Très difficile	Extrêmement difficile	Je ne peux pas le faire, ou c'est un problème très grave	Cochez "ne s'applique pas" si la question ne vous concerne pas ou si vous ne savez pas

15 À votre travail, soulever (des objets ou des personnes)?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

16 Conduire pendant plus d'une demi-heure?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

17 Faire de l'exercice, pratiquer des sports ou d'autres activités récréatives à l'exterieur?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

18 Faire des activités de loisir, comme de l'art ou de l'artisanat, s'adonner à des hobbys, jouer à des jeux de société ou faire de la musique?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

19 Problèmes d'argent à cause d'une perte de revenus liée à votre problème?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

20 Problèmes d'argent à cause du coût des médicaments, des traitements ou des appareils aidant à votre confort ou des frais de gardienne lorsque vous devez aller à vos thérapies ou à vos rendez-vous, à cause de votre problème?

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

SECTION B

À quelle fréquence vivez-vous, *MAINTENANT*, les situations ou les émotions suivantes, À CAUSE DE VOTRE PROBLÈME?
MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS
Jamais	Rarement	Quelquefois	De temps en temps	Souvent	Très souvent	Tout le temps	Cochez "ne s'applique pas" si la situation n'est pas liée à votre problème

21	Je me sens anxieux(se), nerveux(se), inquiet(ète).	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
22	Je me sens triste, déprimé(e) ou de mauvaise humeur.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
23	Je me sens irritable, frustré(e) ou en colère.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
24	J'ai des discussions ou des disputes avec mon (ma) partenaire ou avec d'autres membres de ma famille.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
25	Je me fâche rapidement.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
26	J'ai moins de plaisir dans la vie.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
27	Je me sens mal dans ma peau.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
28	Je me sens moins capable, moins confiant(e) ou moins utile.	1	2	3	4	5	6	7	NE S'APPLIQUE PAS

À quelle fréquence vivez-vous, *MAINTENANT,* les situations ou les émotions suivantes, À CAUSE DE VOTRE PROBLÈME?

MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7
Jamais	Rarement	Quelquefois	De temps en temps	Souvent	Très souvent	Tout le temps

NE S'APPLIQUE PAS

Cochez "ne s'applique pas" si la situation n'est pas liée à votre problème

29 Je sens que je ne peux pas prendre soin de ma famille comme d'habitude.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

30 J'ai peur de ne plus être capable de gagner ma vie à l'avenir.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

31 J'ai de la difficulté à m'endormir.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

32 Je me réveille pendant la nuit parce que je sens de la douleur, des engourdissements ou des picotements.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

33 Je me réveille de bonne heure à cause de la douleur.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

34 Je dors mal.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

35 Je ressens de la douleur ou de l'inconfort (ou des effets secondaires négatifs) pendant ou après les examens, les traitements ou les thérapies pour mon problème au cou ou au membre supérieur.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS

Pour plus d'information sur ce questionnaire, veuillez contacter:

Dr. Susan Stock
Unité de santé au travail et santé environnementale
Direction de la santé publique
75, Port-Royal est, bureau 240
Montréal (QC) H3L 3T1

Téléphone: (514) 858-7510, ext. 263
Télécopieur: (514) 858-5993
Courrier électronique: in44@musicb.mcgill.ca

Annexe C

Résultats des critères d'élimination et de sélection pour les 20 items finals sélectionnés à l'étude anglophone

Item	Applicabilité (% des répondants pour lesquelles l'item a été applicable)		Fiabilité test-re-test (Pearson R)	Corrélation entre l'item et le score moyen du domaine sans l'item	Corrélation entre l'item et l'item global du domaine	Corrélation entre le CRS de l'item et le CRS de l'item global du domaine	Corrélation entre l'item et la question d'évaluation globale par le sujet et le clinicien	Corrélation entre le CRS de l'item et le CRS de la question d'évaluation globale par le sujet et le clinicien
	Hommes	Femmes						
Soins personnels (Q-globale)	100	100	0,79	86	-	-	0,53	0,56
Soins des cheveux	100	100	0,73	0,70	0,73	0,71	0,24	0,51
Atteindre le dos	100	100	0,72	0,77	0,72	0,69	0,26	0,55
Tâches domestiques/ familiales (question globale)	100	100	0,82	0,83	-	-	0,53	0, 56
Ouvrir pots etc.	71,8	89,6	0,82	0,66	0,58	0,63	0,28	0,42
Porter des sacs, caisses	91,5	78,7	0,80	0,73	0,74	0,63	0,38	0,43
Vitesse du travail	100	100	0,76	0,66	0,73	0,60	0,41	0,44
Quantité du travail	100	100	0,83	0,61	0,74	0,75	0,32	0,66
Étirer le bras au travail	91,3	86,7	0,75	0,62	0,53	0,60	0,36	0,54
~Soulever objets hauteur ceinture Soulever objets hauteur épaules Soulever des personnes	78,5 62,9 4,2	79,5 62,5 25,0	0,67	0,73	0,64	0,68	0,22	0,56
Loisirs (Q-globale)	100	100	0,75	0,72	-	-	0,39	0,57
Anxieux, nerveux*	100	100	0,61	0,68	0,60	0,81	0,38	0,37
Triste, déprimé*	100	100	0,58	0,73	0,77	0,79	0,40	0,41
Irritable, frustré...*	100	100	0,72	0,72	0,64	0,69	0,36	0,36
Moins de plaisir*	100	100	0,62	0,73	0,60	0,74	0,28	0,36
Soins de la famille*	100	100	0,69	0,65	0,54	0,44	0,47	0,46
Peur d'incapacité à gagner sa vie*	100	100	0,75	0,57	0,46	0,40	0,21	0,33
Se réveiller la nuit	100	100	0,82	0,77	0,77	0,69	0,31	0,54
Sommeil (Q-globale)	100	100	0,77	0,86	-	-	0,35	0,72
Effets négatifs des examens ou traitements (Q-globale)	100	100	0,74	0,57	-	-	0,49	0,48

~la formulation de la question sur la manutention au travail incluse dans la version finale regroupe des 3 questions testées et elle est donc un peu différente que les versions testées. Pour cette raison elle n'est pas en gras (voir section 3.7)

* formulation avec échelle de réponse «accord-désaccord»

Annexe D

Résultats des critères d'élimination et de sélection pour les 20 items finals sélectionnés à l'étude francophone (formulation originale)

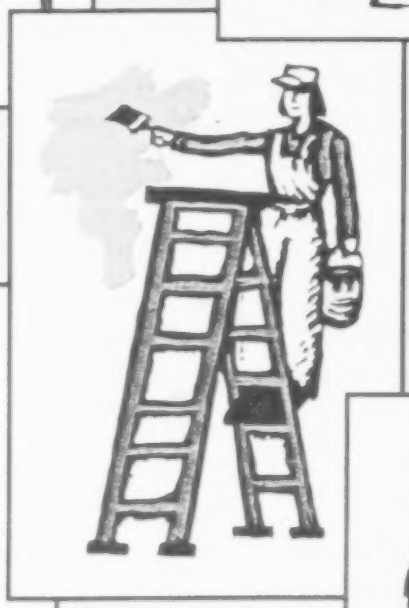
Item	Applicabilité (% des répondants pour lesquelles l'item a été applicable)		Fiabilité test-re-test (Pearson R)	Corrélation entre l'item et le score moyen du domaine sans l'item	Corrélation entre l'item et l'item global du domaine	Corrélation entre le CRS de l'item et le CRS de l'item global du domaine	Corrélation entre l'item et la question d'évaluation globale par le sujet et le clinicien	Corrélation entre le CRS de l'item et le CRS de la question d'évaluation globale par le sujet et le clinicien
	Hommes	Femmes						
Soins personnels (Q-globale)	100	100	0,63	0,89	-	-	0,56	0,54
Soins des cheveux	100	100	0,69	0,74	0,71	0,79	0,56	0,51
Atteindre le dos	100	100	0,78	0,78	0,77	0,80	0,58	0,59
Tâches domestiques/ familiales (question globale)	100	100	0,80	0,83	-	-	0,62	0,55
Ouvrir pots etc.	100	100	0,89	0,65	0,57	0,61	0,49	0,38
Porter des sacs, caisses...	97,6	86,5	0,88	0,82	0,66	0,75	0,64	0,64
Vitesse du travail	95,1	94,2	0,83	0,77	0,78	0,84	0,56	0,74
Quantité du travail	100	98,1	0,81	0,85	0,84	0,87	0,61	0,71
Étirer le bras au travail	87,8	67,3	0,92	0,80	0,66	0,78	0,42	0,79
~Soulever objets hauteur ceinture	87,8	44,2	0,84	0,83	0,85	0,78	0,63	0,72
Soulever objets hauteur épaules	75,6	40,4	0,86	0,87	0,89	0,90	0,63	0,84
Soulever des personnes	2,4	23,1	0,57	0,54	0,59	0,59	0,65	0,56
Loisirs (Q-globale)	97,6	98,1	0,69	0,75	-	-	0,47	0,64
Anxieux, nerveux*	100	100	0,59	0,71	0,61	0,80	0,33	0,58
Triste, déprimé*	100	100	0,55	0,83	0,76	0,80	0,41	0,52
Irritable, frustré...*	100	100	0,75	0,78	0,78	0,81	0,40	0,47
Moins de plaisir*	100	98,1	0,73	0,59	0,56	0,78	0,31	0,58
Soins de la famille*	87,8	98,1	0,59	0,56	0,56	0,67	0,46	0,41
Peur d'incapacité à gagner sa vie*	100	100	0,81	0,71	0,61	0,61	0,21	0,40
Se réveiller la nuit	100	100	0,75	0,71	0,77	0,81	0,26	0,71
Sommeil (Q-globale)	100	100	0,73	0,78	-	-	0,39	0,72
Effets négatifs des examens ou traitements (Q-globale)	100	100	0,50	0,66	-	-	0,36	0,52

~la formulation de la question sur la manutention au travail incluse dans la version finale regroupe des 3 questions testées et elle est donc un peu différente que les versions testées. Pour cette raison elle n'est pas en gras (voir section 3.7)

* formulation avec échelle de réponse «accord-désaccord»

IDVQ-20 **Version 2 (janvier 1998)**

Indice d'Impact des Douleurs du cou et des membres supérieurs sur la Vie Quotidienne



PREVICAP
Centre de recherche
Hôpital Charles-LeMoyne

RÉGIE RÉGIONALE DE LA SANTÉ ET DES SERVICES SOCIAUX DE MONTRÉAL-CENTRE
Direction de la santé publique

INDICE D'IMPACT DES DOULEURS DU COU ET DES MEMBRES SUPÉRIEURS SUR LA VIE QUOTIDIENNE (IDVQ-20)

CONSIGNES POUR REMPLIR CE QUESTIONNAIRE

Par ce questionnaire, nous voulons obtenir des informations sur votre problème au cou ou aux membres supérieurs et votre capacité à effectuer certaines activités. Ceci n'est pas un test et il n'y a pas de bonnes ou de mauvaises réponses. Le temps requis pour compléter le questionnaire est d'environ 10 minutes. Nous vous demandons de répondre à toutes les questions.

Dans la section A, vous allez coter votre capacité à faire diverses activités en vous basant sur votre état au cours de la dernière semaine en encerclant le chiffre le plus approprié. Par exemple, encerclez le chiffre « 1 » si vous n'avez aucune difficulté à faire l'activité ou encerclez le chiffre « 7 » si vous êtes incapable de faire l'activité, ceci dépendamment du niveau de difficulté que vous avez avec l'activité ou la situation énoncée dans la question. Si vous n 'avez pas eu l'opportunité de faire une certaine activité durant la dernière semaine, nous vous demandons d'estimer le mieux possible laquelle des réponses serait la plus précise. Il importe peu que vous utilisiez l'une cu l'autre des deux mains ou bras pour faire l'activité; nous vous demandons de répondre concernant votre capacité d'effectuer l'activité sans vous souciez de la façon dont vous la faites.

Les questions 7 à 10 qui portent sur le travail réfèrent à votre travail habituel. Si vous n'aviez pas d'emploi avant votre blessure, veuillez répondre « ne s'applique pas » à ces questions.

Prenez note que pour la section B (questions 12 à 20), les chiffres de « 1 » à « 7 » mesurent la fréquence à laquelle vous êtes en présence de la situation énoncée dans la question (ex. : « 1 » veut dire « jamais » et « 7 » veut dire « tout le temps »).

Il faut répondre « ne s'applique pas » si une question réfère à une activité que vous ne faites pas habituellement ou encore si la situation décrite dans la question n 'est pas liée à votre problème au cou ou aux membres supérieurs.

SECTION A

Quel degré de difficulté avez-vous ou auriez-vous *MAINTENANT* avec les activités ou les situations suivantes, À CAUSE DE VOTRE PROBLÈME ?
MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS
Aucune difficulté	Un peu difficile	Assez difficile	Moyenne-ment difficile	Très difficile	Extrême-ment difficile	Je ne peux pas le faire	Cochez " ne s'applique pas " si la question ne s'applique pas ou si vous ne savez pas

1	Vous laver, vous habiller, vous arranger ou manger ?	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
2	Vous laver le dessus de la tête ou laver, sécher ou arranger vos cheveux ?	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
3	Atteindre votre dos (par exemple pour laver ou essuyer votre dos, pour sortir votre portefeuille de votre poche arrière ou pour attacher votre soutien-gorge ou une fermeture éclair) ?	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
4	Faire les tâches à la maison ou vous occuper de votre famille ?	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
5	Ouvrir des boîtes de conserve ou des pots, tourner des clés, des robinets ou des poignées de porte ?	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
6	Transporter des sacs d'épicerie, des caisses de bière ou de boissons gazeuses, ou sortir les déchets de la maison ?	1	2	3	4	5	6	7	NE S'APPLIQUE PAS

Quel degré de difficulté avez-vous ou auriez-vous *MAINTENANT* avec les activités ou les situations suivantes, À CAUSE DE VOTRE PROBLÈME ?

MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7	NE S'APPLIQUE PAS
Aucune difficulté	Un peu difficile	Assez difficile	Moyennement difficile	Très difficile	Extrêmement difficile	Je ne peux pas le faire	Cochez " ne s'applique pas " si la question ne s'applique pas ou si vous ne savez pas

7	Faire la vitesse de travail que l'on attend de vous ?							
	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
8	Faire la quantité de travail que l'on attend de vous ?							
	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
9	À votre travail, étirer les bras vers l'avant, vers l'arrière, vers le haut, vers le bas ou sur les côtés, de façon répétitive ou en forçant ?							
	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
10	À votre travail, soulever (des objets ou des personnes) ?							
	1	2	3	4	5	6	7	NE S'APPLIQUE PAS
11	Faire de l'exercice, pratiquer des sports ou faire d'autres activités de loisir comme s'adonner à des hobbys ou jouer des jeux de société ?							
	1	2	3	4	5	6	7	NE S'APPLIQUE PAS

SECTION B

À quelle fréquence vivez-vous, *MAINTENANT*, les situations ou les émotions suivantes, À CAUSE DEVOTRE PROBLÈME ?

MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7
Jamais	Rarement	Quelquefois	De temps en temps	Souvent	Très souvent	Tout le temps

NE S'APPLIQUE PAS

Cochez " ne s'applique pas " si la situation n'est pas liée à votre problème

12 Je me sens anxieux(se), nerveux(se), inquiet(ète).

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

13 Je me sens triste, déprimé(e).

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

14 Je me sens irritable, frustré(e) ou en colère.

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

15 J'ai moins de plaisir dans la vie.

1	2	3	4	S	6	7

NE S'APPLIQUE PAS

16 Je sens que je ne peux pas prendre soin de ma famille comme d'habitude.

1	2	3	4	S	6	7

NE S'APPLIQUE PAS

17 J'ai peur de ne plus être capable de gagner ma vie à l'avenir.

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

À quelle fréquence vivez-vous, *MAINTENANT*, les situations ou les émotions suivantes, À CAUSE DE VOTRE PROBLÈME ?

MAINTENANT veut dire au cours des sept derniers jours.

1	2	3	4	5	6	7
Jamais	Rarement	Quelquefois	De temps en temps	Souvent	Très souvent	Tout le temps

NE S'APPLIQUE PAS

Cochez " ne s'applique pas " si la situation n'est pas liée à votre problème

18 Je me réveille pendant la nuit parce que je sens de la douleur, des engourdissements ou des picotements.

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

19 Je dors mal.

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

20 Je ressens de la douleur ou de l'inconfort (ou des effets secondaires négatifs) pendant ou après les examens, les traitements ou les thérapies pour mon problème au cou ou au membre supérieure.

1	2	3	4	5	6	7

NE S'APPLIQUE PAS

Pour plus d'information sur ce questionnaire,
veuillez contacter :

Dr. Susan Stock
Unité de santé au travail et santé environnementale
Direction de la santé publique
1301, rue Sherbrooke est
Montréal (Québec) H2L 1M3
Canada

Téléphone : (514)528-2400
Télécopieur : (514)528-2459
Courrier électronique : sstock@santepub-mtl.qc.ca

Annexe F

NULI-20 **Version 2 (Jan. 1998)**

Neck and Upper Limb Index

PRÉVICAP
Centre de recherche
Hôpital Charles-LeMoyne

RÉGIE RÉGIONALE DE LA SANTÉ ET DES SERVICES SOCIAUX DE MONTRÉAL-CENTRE
Direction de la santé publique

NECK AND UPPER LIMB INDEX (NULI-20)

INSTRUCTIONS FOR FILLING OUT THIS QUESTIONNAIRE

This questionnaire asks about your neck or upper limb problem and your ability to perform certain activities. This is not a test and there are no right or wrong answers. It will take you approximately 10 minutes to complete. Please answer every question.

In section A, you will rate your ability to do a variety of activities based on your condition in the last week, by circling the most appropriate number. For example, circle number "1" if you have no difficulty at all to perform the task; circle number "7" if you are unable to do the activity, or circle any of the numbers in between depending on the amount of difficulty you have with the activity or situation in the question. If you did not have the opportunity to perform a certain activity in the past week, please make your best estimate on which response would be the most accurate. It doesn't matter which hand or arm you use to perform the activity. Please answer based on your ability regardless of how you perform the task.

Questions 7 to 10 ask about your work and refer to your usual job tasks. If you do not have a job and were not working at the time of your hand or wrist injury, answer "not applicable" to these questions.

Note that for Section B (questions 12 to 20), the numbers "1" to "7" refer to how frequently you experience the situation in the question (e.g. "1" refers to "never"; "7" refers to "all the time").

Please check "not applicable" if the question refers to an activity that you do not normally do or if the situation described in the question is not related to your neck or upper limb.

SECTION A

How much difficulty do you have or would you have *NOW* with the following activities or situations BECAUSE OF YOUR PROBLEM? *NOW* refers to the past 7 days.

1	2	3	4	5	6	7	NOT APPLICABLE
No difficulty at all	A little difficult	Somewhat difficult	Moderately difficult	Very difficult	Extremely difficult	Cannot do	Check "NOT APPLICABLE" if question does not apply to you or you don't know

1	Washing yourself, dressing, grooming or eating?							NOT APPLICABLE
	1	2	3	4	5	6	7	
2	Washing the top of your head or washing, drying or arranging your hair?							NOT APPLICABLE
	1	2	3	4	5	6	7	
3	Reaching behind your back (for example, to wash or dry your back, get your wallet out of your back pocket or do up a bra or back zipper)?							NOT APPLICABLE
	1	2	3	4	5	6	7	
4	Doing household chores or looking after your family?							NOT APPLICABLE
	1	2	3	4	5	6	7	
5	Opening cans, jars or turning keys, taps or door knobs?							NOT APPLICABLE
	1	2	3	4	5	6	7	
6	Carrying grocery bags, cases of beer or soft drinks, or garbage bags?							NOT APPLICABLE
	1	2	3	4	5	6	7	

How much difficulty do you have or would you have *NOW* with the following activities or situations BECAUSE OF YOUR PROBLEM? *NOW* refers to the past 7 days.

1	2	3	4	5	6	7	NOT APPLICABLE
No difficulty at all	A little difficult	Somewhat difficult	Moderately difficult	Very difficult	Extremely difficult	Cannot do	Check "NOT APPLICABLE" if question does not apply to you or you don't know

7 Keeping up with the speed of work expected of you?

1	2	3	4	5	6	7	NOT APPLICABLE

8 Keeping up with the amount of work expected of you?

1	2	3	4	5	6	7	NOT APPLICABLE

9 At work, using your arms to reach forward, backward, up, down, or to the side, repeatedly, or with force?

1	2	3	4	5	6	7	NOT APPLICABLE

10 Lifting (objects or people) at work?

1	2	3	4	5	6	7	NOT APPLICABLE

11 Doing exercise, playing sports or doing other recreational activities such as hobbies or games?

1	2	3	4	5	6	7	NOT APPLICABLE

SECTION B

How often is the following true for you *NOW* BECAUSE OF YOUR PROBLEM?
***NOW* refers to the past 7 days.**

1	2	3	4	5	6	7	NOT APPLICABLE
Never	Rarely	Sometimes	Somewhat often	Often	Very often	All the time	Check "NOT APPLICABLE" if not related to your problem

12 Feeling anxious, nervous, worried or jumpy.

1	2	3	4	5	6	7	NOT APPLICABLE

13 Feeling sad, depressed or moody.

1	2	3	4	5	6	7	NOT APPLICABLE

14 Feeling irritable, frustrated or angry.

1	2	3	4	5	6	7	NOT APPLICABLE

15 Feeling less pleasure in life.

1	2	3	4	5	6	7	NOT APPLICABLE

16 Feeling like I can't take care of my family the same way I used to.

1	2	3	4	5	6	7	NOT APPLICABLE

17 Being afraid that in the future I will be unable to earn my living.

1	2	3	4	5	6	7	NOT APPLICABLE

How often is the following true for you *NOW* BECAUSE OF YOUR PROBLEM?
***NOW* refers to the past 7 days.**

1	2	3	4	5	6	7	NOT APPLICABLE
Never	Rarely	Sometimes	Somewhat often	Often	Very often	All the time	Check "NOT APPLICABLE" if not related to your problem

18 Waking up during the night with pain, tingling or numbness.

1	2	3	4	5	6	7	NOT APPLICABLE

19 Sleeping badly.

1	2	3	4	5	6	7	NOT APPLICABLE

20 Having pain or discomfort from the physical examinations, treatments or therapy for my neck or upper limb problem.

1	2	3	4	5	6	7	NOT APPLICABLE

For more information about this questionnaire, please contact:

Dr. Susan R. Stock
Montréal Department of Public Health
1301, rue Sherbrooke est
Montréal (Québec) H2L 1M3
Canada

Telephone: (5l4) 528-2400
Fax: (514) 528-2459
e-mail: sstock@santepub-mtl.qc.ca

6 champs de recherche en santé et en sécurité du travail

- Accidents
- Bruit et vibrations
- Équipements de protection
- Sécurité des outils, des machines et des procédés industriels
- Substances chimiques et agents biologiques
- Troubles musculo-squelettiques

Nos recherches *travaillent* pour VOUS !

Solidement implanté au Québec depuis 1980, l'Institut de recherche Robert-Sauvé en santé et en sécurité du travail (IRSST) contribue, par la recherche, à la prévention des accidents du travail et des maladies professionnelles et à la réadaptation des travailleurs qui en sont victimes.

POUR EN SAVOIR PLUS...

Visitez notre site Web ! Vous y trouverez une information complète et à jour. De plus, toutes les publications éditées par l'IRSST peuvent être téléchargées gratuitement.

irsst

www.irsst.qc.ca